The Structure of and Incentives from Workers' Compensation Pricing

Karen Roberts
Michigan State University

Chapter 7 (pp. 171-198) in:
Workplace Injuries and Diseases: Prevention and Compensation: Essays in Honor of Terry Thomason
Karen Roberts, John f. Burton, Jr., and Matthew M. Bodah, eds.
Kalamazoo, MI: W.E. Upjohn Institute for Employment Research, 2005
DOI: 10.17848/9781429454919.ch7

Copyright ©2005. W.E. Upjohn Institute for Employment Research. All rights reserved.
Workers’ compensation was originally designed to deal with an aspect of the employment relationship, that of work-related illness and injury. As such, it is viewed as a transaction between two primary stakeholders, the employer and the employee. Examination of the pricing of workers’ compensation requires a shift in orientation. In addition to the employee–employer interaction, the transaction in pricing is also between the employer and the insurer. As is the case in the employee–employer relationship, the two parties have somewhat different interests with respect to the purchase of insurance. Unlike the literature that examines the injury, benefit, and return to work, where the interests of the employer and insurer are treated as coinciding, examination of workers’ compensation pricing is often modeled as a principal–agent problem where information imperfections distort behavioral incentives (Dionne and Harrington 1992). Principal–agent refers to conditions when one party, the principal, either because of a lack of skills or for reasons of cost, contracts with another party, the agent, to perform a particular function. The principal–agent problem arises because the principal is unable, due to costs or other factors, to perfectly monitor the agent. The problem facing the principal is how to structure incentives for the agent to perform the function in the absence of complete monitoring (Sappington 1991).

The purpose of this chapter is to examine how workers’ compensation prices are determined, the incentives embedded in that pricing process, in what ways the employer and insurer perspectives coincide.
and where they diverge, and the policy implications of the different perspectives.

Much of the literature about workers’ compensation pricing is framed in terms of its effect on workplace safety. The theoretical framework is based on commonly made economic assumptions about rationality and utility maximization (Thomason 2003). As is the case with most positive economic models, actor behavior can be understood in terms of a response to benefits and costs (Ehrenberg and Smith 2003). Workers’ compensation is sometimes described as a payroll tax. Employers are seen as trading off between the cost of providing safety and the cost of insurance in the form of workers’ compensation premium. Similar to unemployment insurance, the cost of workers’ compensation insurance is designed to vary with the employer’s use of the insurance, and can thus be seen as lowering the relative cost of safety and providing an incentive to maintain a safer workplace. Employers are therefore seen as engaging in a cost-minimization strategy.

In contrast to unemployment insurance, however, the actor that generates the safety incentive is a profit-maximizing insurer rather than a welfare-maximizing government. Examination of the incentives to insurers shows that the loss-minimizing strategy insurers use is designed to minimize uncertainty around losses rather than minimizing the losses themselves. Economic theory suggests that the workers’ compensation bargain is voluntarily entered into and results in a mutually welfare enhancing arrangement (McCluskey 1998). One question raised in this chapter is whether that is true given the nature of the objective functions of employers and insurers.

The next section of this chapter includes a brief description of how workers’ compensation prices are determined. Subsequent sections discuss the incentives embedded in the price determination process and the research that examines the effectiveness of those behavioral incentives.

**PREMIUM DETERMINATION IN WORKERS’ COMPENSATION**

Premium determination in workers’ compensation is a several-step process, and states vary in the degree to which they regulate the various
steps. Up until the early 1980s, most states used what is called administered pricing to determine workers’ compensation rates. Under administered pricing, states fully regulate prices by issuing basic rates and allowable adjustments that all carriers in the state must use, including the state fund if there is one. Over the past 20 years, many states have deregulated the pricing process to varying degrees (Thomason, Schmidle, and Burton 2001, pp. 39–41). Some states permit carriers to deviate from the basic rates, although these deviations must be applied uniformly across policyholders, and some states permit carriers to set their basic rates independently rather than as a deviation from the state issued rates. Additionally, various states permit carriers to use schedule credits (or debits), which are percentage adjustments to premium that can be applied to individual insurance purchasers and do not have to be offered to all customers. At a minimum, all states require that carriers file their rating plans with the state insurance bureau, but some differ in their policies on whether a carrier must wait for state approval before using their filed rates. For example, file-and-use states permit carriers to use rates as soon as they have been filed with the state bureau, whereas use-and-file states permit carriers to use rates prior to filing. Others require a waiting period to permit the bureau to review the rating plan.

All production activity in the workplace is divided into rating classifications (McGavin 2001). Most states use the classification system devised by the National Council on Compensation Insurance (NCCI), which includes over 600 categories. These categories or class codes are a mix of industry and occupation that are intended to be homogenous with respect to risk. Employers are assigned to these class codes based on their payroll. Class assignment is proportional to payroll distribution.

In each state, a base rate is associated with each of these class codes, expressed in terms of dollars per hundred dollars of payroll. These base rates are developed by rating organizations that collect loss and expense data from insurers, incorporate trend estimates (possibly including a claim adjustment expense factor) and generate rates that are meant to equal expected losses across the state in each of these rating categories (Parry and Math 1993). Workers’ compensation is unique among insurance lines in that all insurers in a state are required to be members of the rating organization (Carroll and Kaestner 1995). There is variation across states in what is included in these base rates. These rates
are referred to as pure premium if they reflect only expected losses. In some states, factors to include loss adjustment expenses are added in to generate what are called loss cost rates. In other states, a profit loading is applied that includes underwriting expenses, commissions and profits, creating what are termed advisory rates (Thomason, Schmidle, and Burton 2001, p. 61). These can be thought of as fully developed manual rates.

In regulated states, the rating organization submits these rates to an insurance bureau or other state agency responsible for insurance industry regulation for approval. This can be a politically charged process as the insurance bureaus are responsible for assuring both statewide insurance industry solvency and employer access to insurance, and may fail to approve rates which the rating organization views as adequate on behalf of its members (Danzon and Harrington 1998). In administered pricing states, once these base rates are approved they become the manual rates all carriers in the state are required to use, with the exception of some states where carriers may apply for approval of an across-the-board deviation. In open competition states, the rating organization is responsible for generating pure premium, loss cost, or advisory rates for insurance commissioner approval. Once approved, these rates become the basis for carriers to independently develop their own manual rates. In open competition states, manual rates are an aspect of price on which carriers can compete for business. Each employer’s manual premium is equal to the manual rate for the class code multiplied by employer payroll in that class code divided by 100.

The next step in the pricing process is experience rating. There are two prevailing descriptions of the purpose of experience rating. One that comes primarily from the insurer perspective is that it is “a procedure that utilizes the individual risk’s past loss experience to forecast future losses. It is an effort to modify the ratemaking process by recognizing an individual risk’s potential for incurring claims” (Parry and Math 1993, p. 658). The other description is framed in terms of providing safety incentives to employers where employers are viewed as responding to an economic incentive to operate a safe workplace: “To the extent that a firm’s own injury experience is reflected in its premium, there is an induced incentive for it to consider investing in safety. If its injuries fall, so will its workers’ compensation premium” (Smith 1993, p. I-152). The distinction between the two descriptions
is that one emphasizes accurate prediction of future losses in order to collect sufficient premium, while the other emphasizes safety-related behavior incentives.

Most states use a standard formula developed by NCCI, although some states permit carriers to develop their own. The basic formula for the experience modification factor is (Chelius and Smith 1993):

\[ X_j = \frac{Ap_j + W_j Ae_j + (1 - W_j) Eej + Bj}{Ei + Bj} , \]

where \( Ap_j \) = firm \( j \)'s actual primary losses; \( Ae_j \) = firm \( j \)'s actual excess losses; \( Eej \) = expected excess losses for a firm of \( j \)'s payroll size in industry \( i \); \( Ei \) = the total expected losses for a firm of \( j \)'s payroll size in industry \( i \) (is equal to the sum of expected primary losses and expected excess losses); and \( W_j \) = is a weighting factor theoretically ranging between 0 and 1 that determines how much weight to give to a firm's own losses (Gillam 1995). The smaller the firm, the smaller \( W_j \) is. \( Bj \) = ballast factor that moves inversely with \( W_j \).

Actual and expected losses are calculated over a three-year period. It is frequently noted that the experience modification formula gives more weight to frequency than severity of injury and that the smaller the firm, the less self-rated it is. The weighting of frequency above severity is accomplished by the separation of losses into actual primary and excess losses. Actual primary losses are equal to actual losses but capped at $5,000. Losses over that amount are included but weighted by firm size. The addition of a ballast factor in both the numerator and denominator reduces the effect of experience rating for smaller firms such that they tend to pay as though they had experienced losses equal to the industry average. Larger firms, on the other hand, face a premium that is adjusted to more closely reflect their true loss experience. The \( W \) factor is termed firm credibility and refers to the degree to which a firm's loss history can be relied upon to predict future losses. Loss histories of larger firms have greater predictive value than smaller firms so \( W \) varies with payroll (Parry and Math 1993). There have been several revisions to the experience rating formula that truncated the true range for \( W \). For example, in Michigan, the possible range for \( W \) is between 0.04 to 0.8 (Compensation Advisory Organization of Michigan 2004). The experience modification factor, \( X_j \), is multiplied by the manual pre-
mium to calculate the firm’s modified, or standard earned, premium (McGavin 2001).

There are several other adjustments that can be made to premium. The most common include premium discounts, schedule credits (or debits), and allowance for deductibles. Actual premium may also be adjusted by dividend payments after the close of the policy period; however, in states that permit ex ante adjustments, evidence suggests that the use of dividends is declining (Yates and Burton 2003, p. 3). Premium discounts are essentially volume discounts and depend on the level of premium generated. In most states, carriers that offer premium discounts are required to apply them uniformly across policyholders.

Schedule credits (debts) are percentage reductions (increases) applied to modified premium. These are considered an underwriting tool in that they allow carriers to adjust premium based on information about expected losses that they do not believe is captured by modified premium. In particular, schedule credits or debits give insurers a mechanism for rewarding employers for improved safety or penalizing for deteriorating safety conditions more rapidly than experience rating permits. Their use also acts as a mechanism for carriers to compete for business on the basis of price.

There are two types of deductible programs, small and large deductible policies. Their use varies by state but nearly all states allow for the use of one or both types of policy. Both types of deductible policy can be seen as a form of “co-insurance” in that they involve explicit sharing of the risk of injury between the insurer and employer (Arrow 1971). Under small deductible policies, employers pay an initial $500 to $2,500 per claim. States vary as to whether they include the deductible amounts in a firm’s loss history. Large deductibles range from between $100,000 to $5 million per accident (Shields, Lu, and Oswalt 1999).

**INCENTIVES TO EMPLOYERS**

The basic framework for understanding incentives to employers begins with the production cost-minimization framework. Injuries impose production costs in the form of lost productivity and perhaps lost human capital investment. Injury costs can be reduced by employer investment in safety, and employers use the efficiency condition such that they
invest in safety up to the point where the marginal benefit equals the marginal cost. The benefit of safety is foregone injury and the cost is an investment cost in equipment, training, or other safety measures. Workers’ compensation benefits are seen as increasing the costs of injuries to employers and thus decreasing the relative cost of safety (Boden 1995). To the extent that employers bear the costs of those workers’ compensation benefits, they will be motivated to invest in safety. The next section discusses the incentives to employers embedded in the various components of workers’ compensation premium determination.

**Manual Rates**

Arguably there are safety incentives for employers embedded in the determination of manual rates (Burton and Chelius 1997). Those rates are determined by the losses experienced by the state’s employers. Better statewide loss histories result in lower manual rates. The use of manual rates as a safety incentive is diluted, however, by free-rider effects, where safer employers subsidize more risky ones who enjoy the lower manual rates without incurring safety costs (Burton and Chelius 1997). Further, it is unlikely that employers understand the possible link between their individual safety records and statewide manual rates (Roberts 2003). There is little research on the effect of manual rates on employer safety behavior. Durbin and Butler (1998) perform national and state-level analyses that separate the effects of manual rates, experience rating, small and large deductibles and OSHA on fatality rates. Using national data, they report that a 10 percent increase in manual rates would lead to a 0.6 percent decrease in fatalities, suggesting that poor understanding and/or issues of cross subsidies undermine potential manual rate safety incentives.

**Firm-Level Experience Rating**

There is a considerable research literature on the effects of experience rating on employer behavior. Existing studies fall into two broad categories: 1) those that use firm size as a proxy for degree of experience rating, and 2) those that examine the effect of experience rating on fatalities.
Firm size–benefit interaction studies

The theory underlying the firm-size studies is based on moral hazard that arises due to information asymmetries in workers’ compensation. Two sorts are considered in the literature: 1) *ex ante* “risk-bearing” moral hazard, where employees reduce their safety efforts because the cost of injury has decreased due to workers’ compensation benefits; and 2) *ex post* “reporting” moral hazard, where employees know the true nature of their injuries, such as whether the injury is truly work-related, how serious it is, and/or whether it genuinely is preventing return to work, but employers do not (Bolduc et al. 2001; Durbin and Butler 1998). As a result, as benefit levels increase, workers’ compensation claims and/or duration will rise due to both types of moral hazard (Krueger 1990). For the firm, therefore, benefit increases raise the marginal savings from injury prevention efforts as long as costs are passed back to firms in a form such as experience rating (Chelius and Smith 1983). The firm size–benefit studies make use of the feature of the experience rating formula that increases the degree of sensitivity to the firm’s own loss experience directly with payroll. Therefore, firm size is treated as a proxy measure for the level of experience rating.

Other than one of the earliest studies by Chelius and Smith (1983) that failed to find support for an experience-rating effect using the benefit–firm size interaction, most studies using this approach do find support for experience rating as a mechanism for reducing claims rates. Ruser (1985) uses data from 25 three-digit manufacturing industries over the 1972–1979 period to test for the benefit-firm size interaction on both frequency and severity of injury. His evidence supports the experience rating effect on the frequency of injury but not for severity. In a later study using establishment data, Ruser (1991) finds strong statistical support for the firm-size effect on injury rates that is robust to several different estimation approaches.

Worrall and Butler (1988) also examine the effect of benefit increases framed as result of two countervailing forces: 1) the employee effect, where higher benefits raise claims costs; and 2) the employer effect, where employers invest in safety to reduce claim costs. They note that if firms were perfectly experience rated, workers would be indifferent between buying their own insurance and a wage adjustment that reflects the compensating wage differential but that reality is plagued with in-
formation failures. Using two-stage least-squares estimation, they estimate the effects wages (result of stage one regression), expected benefit levels, firm size, benefits and firm size squared and the interaction of the two on permanent partial, temporary total, and all indemnity claims.\textsuperscript{2} The sign on the firm size–benefit interaction is negative as they expected, but is not statistically significant in the temporary total benefit claim regression. In a separate study (Butler and Worrall 1988), they examine the effect of experience rating on claim duration using firm size as their indicator of experience rating. Estimating hazard rates, they find that the firm size coefficient is statistically significant, but when used to calculate elasticity, they find that the elasticity of duration with respect to experience rating is substantively small.

Chelius and Smith (1993) take advantage of a unique feature of the Washington State workers’ compensation system, which in addition to conventional experience rating also offers small firms a supplemental experience-rating credit if there are no compensable losses during the rating period. Their prediction is that if experience rating is providing safety behavior incentives, small firms in Washington will have lower injury rates than small firms in other states. The experience rating credit was measured using a dummy variable for the state of Washington. Their results did not support their hypothesis. Arguably, their measure of small firm credit was too imprecise and captured countervailing effects.

**Fatality Rate Studies**

As noted above, one of the difficulties in measuring the effect of experience rating on employer safety behavior is the countervailing effect of benefits on employee claim propensities. The use of fatality rates as an indicator of safety efforts is thought to be free of the effect of both types of moral hazard (Moore and Viscusi 1989). For the most part, studies using this approach find support for a safety effect from experience rating. Moore and Viscusi (1989) use data from the Panel Study of Income Dynamics and the National Institute for Occupational Safety and Health to estimate the effect of workers’ compensation benefits on fatality rates. They use the firm size–benefit generosity measure of experience rating described in the studies above. They find a large negative and statistically significant effect of their measure of experi-
ence rating on fatality rates. Using the same data, they also investigate the extent to which premium costs are passed back to employees in the form of lower wages. They find evidence of positive compensating wage differentials but also that higher wage workers are somewhat more likely to pay for the costs of their benefits. These results indicate that employers enjoy a substantial savings in labor costs by improving safety because much of the cost of safety improvement is financed by employees in the form of lower wages.

Taking advantage of natural experiment presented by the introduction of experience rating in Ontario, Bruce and Atkins (1993) examine the effect of experience rating on fatality rates over a 10-year period. Experience rating is measured by a dummy variable, and industry was the unit of observation. Their results suggest that experience rating in the forestry and construction industries (those subject to experience rating under the Ontario system) led to a permanent (that is, long-run equilibrium) decrease in fatalities in those industries. Further, they found that the magnitude reduced the fatality rate by 20–40 percent.

After providing descriptive evidence that workplace safety was improving despite increases in workers’ compensation costs, Durbin and Butler (1998) examine the effect of several features of insurance pricing on fatalities. They find a large and significant effect of experience rating on the national fatality rate: a 10 percent increase in the cost of experience rating will lead to a 12.3 percent reduction in the number of fatalities.

Potential Weaknesses of the Experience-Rating Studies

There are several potential difficulties with the firm size–benefit interaction approach. One is that the countervailing effect of increased employee claiming may overwhelm empirical evidence of experience-rating effects (Lanoie 1992). The strength of the empirical results suggests that employee efforts may reduce estimates of experience-rating effects, but they do not eliminate them, thus complicating efforts to estimate the magnitude of the effect of experience rating on safety.

A related difficulty is the application of the prevailing assumption about the functioning of reporting moral hazard—that is, that workers’ claiming rates increase with benefit levels. More recent evidence suggests that there is significant underclaiming on the part of injured
workers and that the replacement rate is weakly related to claim behavior (Biddle and Roberts 2003; Shannon and Lowe 2002). This research suggests that there may be countervailing effects other than safety that mitigate against reporting moral hazard.

A third difficulty is determining the true employer safety behaviors in response to experience rating. While the evidence suggests that claims rates do decrease with experience rating, claims rates may decrease due to increased safety investments by employers or by claims management (Thomason 2003). While claims management may entail positive strategies for keeping an injured worker able to work, such as accommodation or rehabilitation, there is also potential for less positive behaviors, such as retaliation for claiming or disputing legitimate claims (Thomason 2003). There is some evidence that experience-rated employers are more likely to dispute claims, though without the ability to measure claim legitimacy, it is not clear that employers are engaging in that behavior (Hyatt and Kralj 1995). In a study of employer response to experience rating, Thomason and Pozzebon (2002) investigate the extent to which employers engage in safety investment and/or positive claims management (accommodation, rehabilitation, and monitoring worker recovery). They find that experience rating leads to more of both safety investment and claims management but that larger firms are more likely to use safety over claims management.

A fourth difficulty arises out of the complexity of the experience-rating formula itself. Arguably, it is sufficiently complex that few employers truly understand the incentives embedded within it (Kralj 1994; Roberts 2003; Spieler 1994). Evidence suggests that the better the firm understands the experience-rating formula, the more likely there will be a safety response, and that understanding increases with firm size (Kralj 1994). This suggests that understanding may be a mediator between experience rating and safety.

Another difficulty is specific to the fatality approach. While the evidence from the fatality rate studies is compelling, the question remains as to whether those results can be safely generalized to the ability of experience rating to reduce the likelihood of nonfatal injuries. These studies appear to implicitly assume that the probability distribution underlying fatal injuries is the same as that for nonfatal. While this may be true in some workplaces, 46.1 percent of lost-time claims in 1999 were sprains, strains, or repetitive motion injuries (U.S. Bureau of La-
abor Statistics 2001), injury types that are unlikely to ever result in death in the extreme. This suggests that results from the fatality studies might be generalizeable to prevention of traumatic injuries such as cuts and lacerations but not to injury rates in general.

Finally, none of these studies actually include a direct measure of the experience-rating modification factor and have to rely on proxies for it, either a firm-size measure or a dummy variable. It is quite possible that those measures are capturing other, unobserved factors that affect injury rates and severity.

**Other Adjustments**

The preponderance of the research on incentives to employers has revolved around experience rating, however; two studies examine the role of deductibles. After a period of rapid growth in workers’ compensation costs during the 1980s, use of large deductible policies increased beginning in the early 1990s (Danzon and Harrington 1998). One study examined the question of whether the use of large deductibles generated incentives for employers to prevent injuries. Taking advantage of a unique research opportunity in Texas, where the use of large deductibles was permitted a year before self-insurance legislation was enacted, Shields et al. (1999) used firm-level data to examine whether permitting deductibles lowered claims rates and claim costs. Their results showed an immediate effect on claim costs, which dropped by 12 percent during the first year after the introduction of large deductible policies. Decreases in claims rates did not appear until the third year. One possible weakness in their study, which they mention but do not address, is that there is likely selection bias due to more risky firms electing to purchase large deductible policies. They also do not discuss what effect the availability of self-insurance a year after the start of their observation period might have had on their sample. However, their results are consistent with those found by Durbin and Butler (1998), who find that small and large deductibles decrease fatality rates by 3.2 percent and 10 percent, respectively.
INSURER PERSPECTIVE

The insurer perspective is commonly modeled as a resource allocation problem, where insurers need to determine the optimal insurance contract. Under the assumption of insurer risk neutrality, the optimal contract is adequate to covering a risk if the policy price is set equal to the expected indemnity plus a loading factor. There are several obstacles to setting this contract, specifically, moral hazard, adverse selection, and regulation (Dionne and Harrington 1992).

Moral hazard, as is the case in its application to claiming behavior, is an artifact of information asymmetry. Again, there are two types of moral hazard, *ex ante* and *ex post*. In this context, opportunities for *ex ante* moral hazard arises because although the insurer can observe that an accident occurred, it can not observe what safety efforts were made to avoid it. The two possible solutions to this are to either provide incomplete coverage against loss or for the insurer to perfectly observe whether the risk is taking care. Only if the latter condition holds is full coverage optimal (Shavell 1979).

*Ex post* moral hazard refers to understanding of the true nature of the accident once it occurs. In most insurance contracts, it is assumed that the insured fully understands the extent of the damage but the insurer does not, and, again, it is costly for the insurer to collect that information. In the context of workers’ compensation, this sort of moral hazard is complicated by the employee’s ability to incompletely disclose the extent of the injury. Adverse selection is also a problem of imperfect information. Resources are misallocated because the insured has no incentive to reveal the true extent of risk and it is expensive for the insurer to observe it (Dionne 1983).

These difficulties with setting the optimal insurance contract provide the logic for three structural features of insurance premium: 1) the creation of risk classes, 2) the use of experience rating, and 3) partial coverage (or deductibles) (Dionne 1983). Because workers’ compensation insurance pricing is regulated, another aspect of the discussion of the optimal insurance contract is how regulation prevents optimality and the unwanted consequences of regulation.
Features of Premium Adjustment

As noted in the description of how premium is determined, manual premium is calculated based on how the firm’s payroll is distributed across rating classes, that is, categories of work activities. The use of rate classifications is one mechanism insurers can use to prevent *ex ante* moral hazard: although the insurer may not be able to perfectly observe safety efforts on the part of the insured, it can approximate risk by classifying the firm’s activities and use market average risk levels for each category to estimate expected losses for the future. Using the argument that to the extent that firms are less able to engage in *ex ante* moral hazard, workplaces will be safer, Lanoie (1992) examined the effect of the number of risk classes on the frequency and severity of injury. The coefficients for the number of rating classes were negative but not statistically significant in the frequency models, and positive and significant in the severity model. These results failed to support his expectation that greater insurer *ex ante* ability to observe true risk, as implied by the ability to assign payroll to detailed rate classes, should reduce the frequency and severity of claims because it is harder for insureds to engage in moral hazard.

Experience rating is used to deal with adverse selection, where firms know how risky their workplace is but insurers do not (Gal and Landsberger 1988). From the insurer perspective, experience rating is functioning well when standard loss ratios (ratio of losses to standard premium) are the same across a recognized risk group (Harrington 1988). In practical terms, this refers to best predicting a risk’s loss for the coming policy period (Sherman 1990). This could be accomplished by manual rates if the risk class were truly homogeneous such that all employers with payroll within each class code maintained equally safe workplaces, but in reality firms are heterogeneous even within rating classes, necessitating additional adjustments.

This perspective on experience rating puts its emphasis on collecting sufficient premium, not on the provision of safety incentives to employers. From the insurer’s perspective, one of the problems with full coverage is that the insured has no motivation to avoid loss and thus will shift all costs of risk to the insurer. For the employer, avoiding loss translates into implementing safety measures, thus incurring costs of safety. Assuming that employers want to minimize costs, they will want
to minimize accidents to the extent that the cost of the losses are passed back to them. However, for the insurer, the key problem is knowing in advance (ex ante) how large the loss will be so they can price accordingly. The size of the prospective loss is not the issue for the insurer, but rather ex ante information about the size of the expected loss. In other words, employers are motivated to minimize loss in order to minimize operating costs, whereas insurers want to minimize uncertainty about the size of the prospective loss.

There are two broad threads in the insurer perspective literature on experience rating: 1) the development of theoretical models that examine the assumptions and conditions under which experience rating generates an optimal contract, and 2) empirical examination of the ways in which the application of experience rating is distorted either by features of the formula itself or through regulation.

The theoretical discussion notes that the way experience rating is incorporated into insurance contracts is based on the Law of the Iterated Logarithm. When applied to the insurance contract, this law provides an understanding of insurance as a set of multi-period contracts. Repeated observation over the length of the relationship approximates perfect observation, thus eliminating the opportunity for moral hazard (Dionne 1983; Dionne and Harrington 1992).

One model suggests that repeated contracts are not necessary for optimality. According to this model, at the end of each period, the insurer offers a contract where premium paid at the end of the period depends on the number of claims submitted until then (Gal and Landsberger 1988). Insureds know that the premium will depend on the number of claims submitted by the end of the period and so recognize there is some uncertainty about the premium since they do not know in advance what their claims rate will be. Thus, uncertainty is shared; but, by paying premium that reflects actual claims made in the last period at the start of the next period, the contract allows for the eventual coverage of losses. The model requires the insured to renew the contract and that losses be accurately reported. The former requirement is probably reasonably accurate for workers’ compensation where insurance is compulsory, but problems with injury underreporting in workers’ compensation suggest that the latter assumption cannot be made safely.

The less theoretical work examining the effect of experience rating examines the accuracy of the formula in adjusting premium so that poli-
cyholders with equal risk produce equal expected loss ratios (Parry and Math 1993). Several criticisms have been leveled against the standard (NCCI) experience rating formula, including 1) a lack of precision in incorporating the tendency for larger firms to have lower loss ratios that result in insufficient manual rates, potentially resulting in availability problems for small firms (Harrington 1988); 2) a tendency to obscure commonality within a risk class (the argument here seems to be that experience rating permits overly heterogeneous risks to be included in one rating class through the process of off-balancing); and 3) an inability to adequately handle shifts in firm ownership that change the underlying riskiness (Parry and Math 1993). Revisions of the standard experience-rating formula in 1993 and again in 1997–1998 along with the use of additional adjustments such as schedule credits, were designed to ameliorate these problems (Compensation Advisory Organization of Michigan 2004; Gillam 1995). In addition, the most recent revision of the formula reduced the effect of medical-only claims on the experience modification factor (Compensation Advisory Organization of Michigan 1998). Again, the reason for this change can be interpreted in two ways. For the employer, this revision arguably provides an incentive to either keep injured workers at work or return them before the waiting period expires. However, for the carrier, a medical-only claim has more predictable future losses, that is, less uncertainty.

THE EFFECTS OF REGULATION

Pricing in the insurance industry is heavily regulated, and because coverage is required for nearly all employers, workers' compensation pricing is more regulated than most other lines (Carroll 1993). Over the past 20 years, however, pricing in workers' compensation has undergone significant reform so that 37 states have enacted some form of open competition since the early 1980s (Thomason, Schmidle, and Burton 2001 p. 42). However, even under open competition, some regulation in the form of rate approval is often required (Hunt, Krueger, and Burton 1988; Thomason, Schmidle, and Burton 2001). The mandate of the state insurance bureaus is to assure insurer solvency, rate affordability, rate fairness, and universal access (Carroll and Kaestner 1995). Typically, workplace safety is not part of the insurance bureau portfo-
lio unless the claims portion of the state regulatory structure is housed there. One of the questions investigated with respect to the effects of regulation in workers’ compensation is the extent to which it creates cross-subsidies that damage both employer and insurer incentives.

One reason that regulation can be thought of as a barrier to the optimal insurance contract is due to the regulatory objectives, which include both a healthy insurance market and affordable rates consistent with a positive economic development climate, and thus require a balancing of competing interests. If, through regulation, insurance bureaus set rates and permit adjustments that are actuarially fair, then the optimal contract will be possible (Schmidle 1995). However, political pressures from employers to lower rates or from insurers to raise them can distort contract incentives and threaten either of the core regulatory objectives.

Typically, there are two primary forms of rate regulation: 1) limit the percentage experience modification factors applied to class rates; and 2) restrict the class rates themselves (Harrington and Danzon 2000). As described earlier, in every state, the rating organization issues basic manual rates (e.g., pure premium, loss cost, or advisory rates). In theory, the rating organization is submitting rates that reflect state average expected losses in each rating class and are actuarially fair. In regulated states, these rates go to the insurance commissioner for approval before being issued for carrier use. In open competition states, carriers use these as guidelines but file their own rates. In principle, each carrier’s rates reflect its own underwriting experience, cost structure, and marketing strategy.

That regulation is beneficial rests on two arguments. One is that government may possess information either not available to the parties or that would be inefficiently costly for each party to uncover for itself, such as insurer solvency or workplace safety. Because of the cost of acquiring that information, data collection and distribution in the form of rates is treated as a public good. The second is a welfare argument that insurance access is important, even if its guarantee involves some inefficiencies (Carroll and Kaestner 1995).

There are several reasons, however, why regulation may be undesirable. Because rate regulation is partially a political process, there may be pressure to suppress overall rates in the name of economic competitiveness. Further, some employers may be more able to influence
the regulatory process leading to the suppression of some rates and a possible compensating inflation of others (Carroll and Kaestner 1995). In addition, it is sometimes argued that the insurance bureau becomes “captive” of the insurance industry and excessively raises rates (Harrington 1984).

A primary criticism of regulation is its potential for generating inequitable and welfare reducing cross-subsidies (Danzon and Harrington 1998). The research on possible cross-subsidies argues that regulation results in subsidies from safer to more risky employers, often in the form of cost shifting from the residual to the voluntary market. The central premise to the cross-subsidy discussion is that if rate regulation results in rates that are insufficient relative to claims costs, distortions will occur that will in fact lead to even greater cost growth (Danzon and Harrington 1998).

Noting that the insurance market is segmented into two portions, the voluntary market and the residual market, one premise that underpins this discussion is that the larger the residual market share, the more troubled the state’s insurance market. All states except those that are exclusive state funds have a residual, or assigned risk, market. Risks that fail to meet the underwriting standards of the voluntary market, are uniquely dangerous, have no loss experience, have had a recent significant loss, or have a loss experience that lacks credibility are typically insured in the residual market. To some extent, the boundary between the two markets is fluid. To the extent that rates are suppressed below actual expected losses in the voluntary market, insurers will be unwilling to insure some of the less safe firms at the margin, forcing them into the residual market. In most states, a cohort of servicing carriers writes the policies and pays the claim but does not bear the underwriting risk in the residual market. Rather, losses incurred in the residual market are shared out among the voluntary market carriers according to each carrier’s share of the state’s premium.

It is this sharing out process that creates the most common form of cross-subsidy (Danzon and Harrington 1998). To the extent that regulation suppresses rates in the residual market in order to assure access to coverage for those firms in the assigned risk pool, aggregate residual market premium will not cover the actual losses along with expenses and fees, generating an excess loss that needs to be capitalized into
Regulators have several different tools for manipulating rates in the residual market, including price differential programs (employers charged a fixed percentage higher than comparable risks in the voluntary market); risk-adjustment programs (ARAP, or Assigned Risk Adjustment Programs, where employers in the residual market are surcharged up to 49 percent of base premiums); restrict residual market access (where employers are denied residual market coverage if they have refused any voluntary market carrier’s legitimate rating plan); removal of premium discount plans (employers in the residual market cannot receive premium discounts even if they would qualify in the voluntary market); and take-out credit programs (where voluntary insurers receive credits for offering residual market firms the same coverage for at least one policy period). Each of these tools acts as a ceiling on residual market rates (Kwon and Grace 1996). Additional variations in the structures of residual market prices include the use of loss-sensitive rating plans that enhance the effect of the individual employer’s loss history, and additional surcharges on employers in the pool (Thomason, Schmidle, and Burton 2001, p. 338).

As residual market losses are passed back to the voluntary market, costs in the voluntary market are expected to increase, and safety in the workplace is expected to deteriorate for several reasons. One is that as premium in the voluntary market increases due to the residual market losses, the larger and safer firms in the voluntary market will self-insure, leaving the less safe in the voluntary market where costs will increase (Danzon and Harrington 1998). This can become a vicious circle if regulation of voluntary market rates does not adjust to take this into account, leading to the more marginal employers in the voluntary market being forced into the residual market. During the 1980s, the markets in several states, notably Maine, Rhode Island, and Louisiana, were caught in this cycle that ended with an inordinately large share of the workers’ compensation market in the assigned risk pool, creating true crises for workers’ compensation in their states (Thomason, Schmidle, and Burton 2001). According to this view of market dynamics, both the size of the gap between premium and expected losses and uncertainty about what that will be lead to a reduction in supply of insurance in the voluntary market, leading to an availability problem and higher rates.
This cycle not only creates availability problems but also dilutes safety incentives. Because regulated premium in the residual market does not correspond to expected losses, those in the residual market will lose the incentive to invest in safety, increasing their loss rates. Again, because these costs are passed to the voluntary market, the cycle described above will be triggered.

Several forms of cross-subsidies are possible, but in all cases, safer employers are subsidizing riskier ones. If all rate class prices are suppressed, those employers in the residual market will be subsidized by those in the same rate class in the voluntary market, as described above. This subsidy will worsen as the higher risk employers are forced into the residual market and the lowest risk employers go to self-insurance, leaving the middle risk employers to subsidize the riskiest. If regulation varies across rate classes, those with payroll in the suppressed rate classes will be subsidized by those in the other classes, whose rates will have increased so that total premium will, on average, cover total losses (Danzon and Harrington 1998).

Because cross-subsidization is difficult to observe, the empirical work examines the broader predictions of these models. Danzon and Harrington (1988) examine the effect of rate regulation on total and indemnity loss growth using three measures of regulation: 1) lagged residual market share, 2) lagged filed to approved rate ratio, and 3) lagged underwriting margin for the filed to approved rate ratio. Their results show a positive and statistically significant relationship between residual market share and loss growth but do not support such a relationship for the other, more direct, measures of regulation. However, using a different specification, they do find significant effects of the lagged filed to approved rate ratio on total loss growth (Harrington and Danzon 2000).

While these results do suggest that regulation adversely affects loss growth, they do not necessarily indicate the presence of cross-subsidies. As an alternative test for cross-subsidies, Danzon and Harrington (1998) develop a model of political influence, which predicts that if regulation of rates is the outgrowth of political influence, patterns of rate suppression should persist over time. Using data from 150 class codes in seven states, the null hypothesis that they test empirically is that fluctuation of manual rates should be random over time. Rejection of the null hypothesis would be viewed as evidence of political influ-
ence in the regulatory process. They examine several potential sources of political influence. They find support for persistent cross-subsidies along three dimensions: 1) between rating classes (larger classes are subsidized), 2) between low- and high-risk employers (high-risk, low-wage employers are subsidized), and 3) between policyholders and insurers (policyholders are subsidized).

Kwon and Grace (1996) examine the magnitude of the cross-subsidy of low risk to high risk employers by examining the share of residual market assessments carriers can pass on to voluntary market policyholders. They find that operating losses in the residual market do raise premium in the voluntary market, but that only about 27 percent of the residual market assessment on carriers is passed on in voluntary premium. They also find that few of the policy measures designed to reduce the size of the residual market are effective. Thomason, Schmidle, and Burton (2001) cite the importance of the Danzon–Harrington perspective in that it examines the distributional consequences of regulation and not just the effect on price levels. However, they also cite the weaknesses of the measures of regulation, in particular, the lack of differentiation among the different regulatory regimes. In addition, their empirical test finds weak support for increased injury rates under administered pricing (the most regulated scheme) and a lack of support for a relationship between deregulation and injury (p. 264). Both the Kwon and Grace study and those by Danzon and Harrington cite the health of the insurance market as the policy focus, but only Danzon and Harrington also mention safety implications.

In their study of the effect of differing insurance arrangements on various workers’ compensation outcomes, Thomason, Schmidle, and Burton (2001, p. 245) find support for full deregulation as being more efficient at reducing employer costs. However, they qualify for this conclusion because of the difficulty of accurately characterizing prederegulation conditions and with evaluating the sustainability of the effects of deregulation.

CONCLUSION

This chapter provided a description of how workers’ compensation premium is determined and then examined the research literature on pric-
ing from two perspectives, that of the employer and that of the insurer. There are some differences between these two perspectives that may have welfare implications. What becomes evident in the research is that the focus in the employer-based literature is on safety incentives embedded in workers’ compensation premium, whereas the emphasis in the insurer-based literature is on generating sufficient premium to cover losses. The significance of this difference is that while the purpose of the premium setting mechanism may be seen as to promote safety from one point of view, from the other it is not to minimize injury but rather to estimate it accurately.

In an unfettered market, this difference in objective functions might not matter: the invisible hand would keep costs low, promote safety, and protect insurer solvency. But the evidence on the safety effects of experience rating (the feature of pricing that has received the most empirical attention) is mixed and potentially quite flawed by the absence of a direct measure of experience rating. Further, the only research that directly examines whether employers understand the incentives workers’ compensation is designed to promote suggests that the formula is not universally well understood (Kralj 1994).

The weaknesses in the experience-rating literature best frame the directions for future research. One particularly weak aspect of this research is the lack of studies that actually include the experience-rating modification factor as the measure of experience rating. Because the experience modification is observable, this seems a feasible research direction. A second research area that would benefit from further investigation is the role that employer understanding of the structure of their insurance premiums plays in safety behavior. With the exception of the few studies cited above, little survey research of employer understanding of or response to safety incentives in workers’ compensation has been done.

The important question that this raises is whether it is sound policy to rely on market forces to the extent that we do to promote safety at work. It appears that U.S. workplaces are becoming safer (Durbin and Butler 1998), with the number of nonfatal illnesses and injuries per 10,000 full-time workers dropping from 304.7 in 1992 to 169.1 in 2001 (U.S. Bureau of Labor Statistics 2004). However, the extent to which improved safety can be attributed to employers responding to the economic incentives from prices has not been established. It may be more
plausible that employers provide a safe workplace for other reasons, such as a belief that it is a good investment (Spieler 1994). Failure to understand the underlying motivational mechanisms will limit effective policy development should the trend toward lower injury rates change.

Notes

1. The discussion in this chapter describes what is referred to as the voluntary market, where employers purchase workers' compensation insurance from insurance carriers. Depending on the year, approximately 20–25 percent of workers' compensation insurance benefits are paid by self-insured, that is, employers who choose to pay workers' compensation benefits directly (Mont et al. 2001, p. 9). The principal-agent framework would not apply to the self-insured portion of the workers' compensation market.

2. Two-stage least-square estimation is an econometric method used to address possible problems with endogeneity—where what is being treated as an independent variable is correlated with the residual and thus is not truly exogenous.

References


Workplace Injuries and Diseases
Prevention and Compensation

Essays in Honor of Terry Thomason

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

2005

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