SOME BENEFIT CONSIDERATIONS IN WORKERS' COMPENSATION

(Prepared for the Royal Commission on Workers' Compensation in British Columbia)

by

Morley Gunderson* and Douglas Hyatt**

(June 1998)

* Professor, Centre for Industrial Relations and Department of Economics, and Member Institute for Policy Analysis, University of Toronto.

** Associate Professor, Scarborough College and Faculty of Management, Member Institute for Policy Analysis, and Scientist Institute for Work and Health, University of Toronto.
SOME BENEFIT CONSIDERATIONS IN WORKERS' COMPENSATION

The benefit level is clearly one of the most high profile design features of workers’ compensation systems, in part because of the central role they play in determining the adequacy of the compensation received by injured workers and because of its potential role in influencing both the likelihood and duration of claims. Such effects in turn can have important implications for regulatory requirements with respect to such factors as vocational rehabilitation, return to work requirements, reasonable accommodation requirements, and provisions that deem workers to be able to return to work and earn a specified level of income consistent with suitable and available jobs. Benefit structures have important ramifications for the costs of the system, both transfer costs as well as real resource costs associated with administering the system. They also have implications for flexibility and adaptability, transparency, and the attitudes of the stakeholders.

Not surprisingly, given the multiple objectives to be served by this seemingly simple policy instrument, numerous trade-offs are involved. Designing the benefit structure to achieve one objective, inevitably will conflict with others. The purpose of this report is to elucidate those trade-offs so as to facilitate informed policy choices, and to highlight the implications of the different options. The analysis focuses on compensation for permanent partial disabilities, since that is the area where the trade-offs are best illustrated.

The report begins by outlining the main generic options, namely scheduled benefit payments ("meat chart") versus wage loss systems. The prevalence of these different options across the various Canadian jurisdictions is then outlined. The incentive effects of the different options are discussed, focusing on the incentives to take adequate precautions against risk, to make claims and to return to
work. Evidence on those impacts is presented. Regulatory incentives are also discussed, with respect to such factors as vocational rehabilitation, return to work requirements, reasonable accommodation requirements, and deeming provisions. The different options are then analyzed with respect the criteria of adequacy, equity and efficiency. Ancillary issues relevant to both options are whether the benefit should be paid on the basis of gross or net earnings and the extent to which benefits from non-workers’ compensation sources should be stacked or integrated with the workers’ compensation benefit. Both of these issues are fully explored in the Workers’ Compensation Board of British Columbia’s briefing papers entitled, “Compensation Rate” (April 4, 1997), “Average Earnings” (April 10, 1997) and “Benefit Stacking and Integration” (May 21, 1997). We offer only some brief comments on these issues in this paper.

As indicated, the emphasis is on elucidating the trade-offs that are involved with respect to evaluating the ability of the different policy options to achieve the multiple objectives of the program. The intent is to identify elements of a "win-win" strategy that will enable the system to deliver its outcomes in an effective fashion. Doing otherwise, will jeopardize the long-run viability of the system either through a loss of stakeholder commitment to the workers’ compensation system, or through escalations in cost. An exclusive focus on cost containment runs the risk of severe inequities that can jeopardize the acceptability of the system, and lead to cost shifting to other systems. Ignoring costs runs the risk of inducing more draconian reforms, especially if the costs jeopardize investment and the jobs associated with that investment. Clearly all parties -- employees, employers and governments -- have a stake in this crucial design feature of workers' compensation systems.
For expository purposes, it is useful to distinguish between two main generic policy options with respect to compensation for permanent partial disabilities: the disability rating schedule and the wage loss system. Combinations of the two are also possible.

Under the *disability rating schedule* ("meat chart") payments are designed to approximate the percentage of permanent disability experienced by the worker based on the workers residual impairment. This disability rating or percent is then multiplied by the workers' pre-injury earnings to determine the amount of compensation. A disability rating of 35% implies that the worker permanently lost 35% of their earnings capacity, and hence they would be compensated by 35% of their earnings. This rating would apply whether the injury actually affected their earnings by 35% or by a larger or smaller amount. In effect, the payment is not conditional upon their actual earnings or work behaviour after the injury. In the jargon of economists, this payment is an unconditional lump-sum transfer or a demogrant (a grant given to the demographic group of workers with a permanent partial disability).

Under the *wage loss* system workers are compensated for a portion of their lost earnings. That portion typically ranges from 75% to 90% of their net earnings loss. Net earnings are typically gross earnings less income taxes and contributions to the Canada Pension Plan and Employment Insurance. The percent may also be based on gross earnings, in which case it tends to be at the lower end of the spectrum, such as 75% of gross earnings. The *actual* income replacement rates will be lower for persons with high earnings since they are generally subject to a maximum. The benefits are also generally subject to a minimum so that low income persons will receive 100% of their earnings loss if they are below the minimum. For these reasons the actual replacement rates will differ from the usual legislated rates, with the average actual rate being lower (shown subsequently) since
considerably more persons are affected by the maximum than by the minimum. The real after-tax replacement rate will be higher than the nominal rate because workers' compensation payments are not taxable and the recipient who does not return to work does not incur work related expenses. On the other hand, the real replacement rate may be lower because it typically does not include fringe benefits.

Earnings losses are estimated on the basis of the difference between the what the person could be expected to earn if their pre-injury wage continued, compared to their actual wage after the injury. Periodic re-assessments are made (e.g., after one year, three years and six years in Ontario) to ensure that compensation benefits are equivalent to the difference between pre- and post-accident earnings. If the Board deems that their actual post-injury wage is not representative of what the claimant could earn, they may be deemed to have a higher wage based on "suitable and available" employment.

In addition to these deeming provisions, wage loss systems often provide for other regulatory requirements designed to facilitate and encourage the return to work of injured workers\(^1\). Vocational rehabilitation is emphasized and co-operation in this area is an important requirement for receiving the indemnity benefit. Employers are required to return the injured worker preferably to their pre-injury job, or to a comparable job\(^2\) if the pre-injury one is not available, or to the next most suitable job if a comparable job is not available. Employers are also required to reasonably accommodate the return to work of injured workers, with reasonable accommodation generally being interpreted as up to the point of "undue hardship." In Ontario the guidelines of the Ontario Human Rights Commission are used to interpret undue hardship. Those guidelines state that for the costs of accommodation to be considered as constituting undue hardship, they must be quantifiable, shown to be related to the accommodation, and so substantial that they would substantially affect the essential nature of the enterprise, or so significant that they would substantially affect the viability of the enterprise. The costs must be firmly documented, and they must be spread over the whole organization, not just the unit doing the accommodation. The employer must make every effort to mitigate those costs through outside support or subsidies, and must be prepared to borrow if necessary.

\(^1\) The regulatory provisions described here pertain to Ontario which has a wage loss system.
\(^2\) A comparable job is interpreted as one with a high degree of similarity in terms of the demands, rewards, status and opportunities, with comparable wages taken to be at least 90% of pre-accident wages.
Dual award systems such as in Ontario can involve compensation for this future economic loss, as well as a lump-sum amount for non-economic loss, based on the degree of disability and independent of the person's earnings. That degree of disability is determined in the same fashion as the "meat chart" with the percent being multiplied by a fixed amount (e.g., around $50,000 in Ontario) to yield a lump-sum amount independent of the person's earnings. That amount is typically adjusted for the age of the worker. For example, in Ontario, it is adjusted downwards by about $1,000 for every year that the worker is older than age 45 (to a minimum of around $28,000) and upwards by $1000 for every year the person is under the age of 45 (to a maximum of around $74,000). Workers who are permanently injured but experience no wage loss would receive only the non-economic loss benefit.

Permanent total disability benefits are paid to workers whose injury leaves them totally disabled and unable to return to work. In systems that use the disability rating schedule ("meat chart") the disability rating is in effect 100% so that the worker receives the equivalent of a temporary total disability (or short-term disability) benefit. In wage loss systems, the post-injury wage effectively is set to zero to so that the wage loss is total.

COMPARISONS ACROSS JURISDICTIONS

Disability rating schedule ("meat chart") systems for permanent partial disabilities exist in Alberta, Prince Edward Island and the Northwest Territories. B.C. uses a hybrid approach under which the worker receives the greater of the functional disability award or the wage loss award. The pension is paid for life. However, for those workers who received a wage loss pension, their benefit is reduced to the functional disability award amount after age 65. The other jurisdictions utilize a wage loss system, or a dual award system. The trend tends to be away from the disability rating
schedule and to a wage loss or dual award system. No system has retreated from a wage loss approach back to a disability rating schedule.

For temporary disabilities, all Canadian jurisdictions compensate on the basis of replacing a portion (usually 75 to 90%) of lost earnings (gross or net, depending on the jurisdiction). The period of temporary disability usually lasts up to the point of maximum medical improvement, after which benefit payments are discontinued or, if relevant, converted to a permanent disability award.

Jurisdictions that use the wage loss system for permanent partial disabilities tend to use the same earnings replacement rates as they use to calculate temporary disability benefits. Alternatively stated, they continue the earnings replacement rate that exists in their temporary disability system for claims that are converted to permanent disability claims.

Table 1 gives the earnings replacement rates for temporary disabilities across the various jurisdictions in Canada. These rates would also apply to permanent partial disability payments in jurisdictions that have wage loss systems; that is, all but, Alberta, Prince Edward Island and the North West Territories, which use disability rating schedules exclusively for determining permanent partial disability benefits, and for functional disability awards in British Columbia.

As indicated in Table 1, the most common income replacement rate is 90% of net earnings loss, although it is 85% in New Brunswick and Nova Scotia (corresponding to the rates that apply after 39 and 26 weeks respectively under temporary disability). The replacement rate in B.C. and Yukon is 75% applied to gross earnings. Maximums in the neighborhood of $500 to $700 per week are common. The maximums of $500 to $700 per week would translate into annual maximums respectively of $26,000 to $36,400 based on 52 weeks. The maximums are close to the average weekly earnings of approximately $600 per week for all workers in Canada as of January 1997.
(Statistics Canada, 1997). The fact that persons above the maximum would not be receiving the full, say, 90% replacement rate, and that benefits may be reduced slightly because of waiting periods, implies that the actual replacement rate will be less than the potential replacement rate of, say, 90%. Thomason and Hyatt (1997, p. 244), for example, indicate that the average actual replacement rate in Canada was approximately 68% of wages in 1993. They have changed markedly over time, however, declining from 65% in 1961, to 58% in 1971, increasing to a high of 70% in 1986, and declining slightly to 68% by 1993.

THEORETICALLY EXPECTED INCENTIVE EFFECTS OF BENEFITS

The main incentive effects of interest under workers compensation pertain to the incentive to take precautions and reduce risks and accidents, the incentive to file claims, and the incentive to return to work. These incentive effects can apply to both employees and employers. In this section, the theoretically expected relationship between workers' compensation benefits and these two outcomes are analyzed. The next section discusses the empirical evidence on these effects.

Theoretically Expected Effect on Real Accident Rates

With respect to accident prevention and risk reduction, higher benefit compensation under workers' compensation is expected to have two main effects: a real accident effect; and a reporting effect. If the compensation benefits are high, employees may have less incentive to take precautions against risks since the monetary consequences of the risk are offset in part by the compensation.
Working in the opposite direction, *employers* would have a monetary incentive to reduce risks, but only if the higher benefit payouts translated into higher costs for employers (e.g., through experience rating). In essence, the ultimate impact of higher benefit rates on accident rates is theoretically indeterminate; it depends upon whether the employee effect to take fewer precautions dominates the employer effect to take more precautions.

This potential impact on real accident rates is often termed the moral hazard effect. This is a term borrowed from the insurance literature, referring to the reduced incentives to take precautions against risk as the risk becomes fully insured. The moral hazard problem can become potentially acute as the income loss becomes fully replaced by the insurance. It is the reason that insurance rates increase substantially if persons try to buy full insurance against risk; conversely, it is the reason that insurance rates can drop substantially if there are large deductibles so that the insured person also bears substantial risk. The moral hazard problem obviously is greater when individuals can change their behaviour to affect the risk.

In general, there is little controversy when the principle of moral hazard is used with respect to the insurance of property. Few would argue that allowing sole-proprietors to insure their plants and equipment for 200% of their value would not lead to more fires among the insured buildings. Most building owners would still take normal precautions, but some may reduce their precautions and a few may even have their buildings burned.

However, when this principle is applied to employee precautions under workers' compensation it generates considerable controversy. This is so because it conjures up the image of people deliberately injuring themselves or taking risks just to get the compensation. Such an implication can be particularly offensive to persons who are accidentally injured at the workplace and
who are now also blamed for bringing it upon themselves -- effectively adding "insult to injury." Certainly it would be a rare occurrence for an individual to deliberately injure themselves to obtain compensation. Furthermore, most workers would not likely alter their precautions or risk taking in response to compensation, in part because it is usually not possible to control against severe or ultimate "downside" risk -- fatalities can obviously occur. However, all that is required for benefit rates to have an effect on accidents is for some workers to make marginal adjustments to their risk taking, even if the vast majority do not alter their behaviour.

Certainly, there is likely to be less controversy over whether profit-maximizing employers would alter their precautions, enforce them more stringently, or manage workers’ compensation claims more aggressively, if greater benefit payouts affected their costs relative to those of their competitors (as experience rating would ensure).

**Theoretically Expected Effect on Reporting of Claims**

Higher compensation benefit rates can also affect the incentive to report accidents and contest claims, given that more is now "at stake". For employees this can occur in various forms. There is a greater monetary incentive to now report apparently minor injuries rather than to "work through" them. There may be a greater incentive to report "off-the-job" injuries as now being "on-the-job" injuries since the later can be compensated by workers' compensation. There may also be a greater incentive to report problems that are not injury related as now being injury related problems given the compensation may only be available to the worker for injury related ones.

These reporting incentive effects are presumed to be more likely to occur for hard-to-detect conditions such as soft-tissue sprains and strains, as opposed to more observable or easily diagnosed
physical injuries. The reporting incentive effect is also likely to be stronger if there is retrenchment in other sources of income support such as unemployment insurance, private disability, Canada Pension Plan disability or social assistance.

While these reporting or claims effects are often suspected of being fraudulent, this is unlikely to be the case in an administrative system in which there is considerable monitoring of claims in pay. The increased reporting of minor claims, for example, could well reflect the reporting of claims that merit reporting; it was simply not worthwhile, if benefits were small. The off-the-job injuries could very well be ones that were induced by job related factors; it was simply not worthwhile to contest that issue if benefits are small. The same applies to problems that are not specifically injury related, but that could reflect the cumulation of work-related causes. This is obviously a "grey" area, and normative judgement calls are very difficult to make on the "appropriate" shade of grey. Changes in the benefit rates, however, are likely to provide an incentive for employees to try to change where they are on that spectrum.

For employers the higher compensation benefits may induce them to contest claims more strenuously if the higher benefits would feed back to their costs, as would be the case for example with experience rating or other penalties. They may even seek to reduce the reporting of accidents if such reporting would otherwise increase their claims costs.

As with the true accident effect, the impact of higher benefits on the reporting and claims effect is theoretically indeterminate. Employees have a greater incentive to enhance claims even if the accident were minor, or occurred off the job, or was not related to the injury, while employers have a greater incentive to contest and reduce claims when benefits are high. This also highlights the reality that accidents do not equal claims.
Compensating Wage Premiums

The previously discussed effects are complicated by the possible existence of compensating wage premiums for risk. If compensating wage premiums fully compensate for risk then changes in the benefit rates would have no effect on accident prevention or reporting; changes in the benefit rate would simply be offset by changes in the compensating wage premium that is paid for uninsured risk.

Compensating wage premiums, to the extent that they prevail, would generally compensate only for uninsured risk. In the extreme, if the benefit rate fully offset the disutility of the injury (an extreme assumption since benefits are generally designed to compensate for lost income) then workers would face no uninsured risk and the compensating wage premium would be zero. Increasing the benefit rate from one that partially compensated for uninsured risk towards one that fully compensated for uninsured risk would therefore reduce the compensating wage premium for uncompensated risk, with the compensating wage premium being zero in the extreme case where the benefit rate fully insured against the risk. In general, the higher the benefit rate, the lower the uninsured risk and hence the lower the compensating wage premium that must be paid.

If compensating wage premiums fully compensated for uninsured risk, then increases in benefit rates would simply reduce the uncompensated risk and the associated compensating wage premium. Workers would have higher ex post compensation if they were injured, but they would have lower ex ante compensating wage premiums in their jobs. In essence, they would be indifferent between a scenario of high benefit rates and low compensating wage premiums, and one of low benefit rates and high compensating wage premiums. In such circumstances, they have no incentive to alter their precautions to change risk -- they are compensated for the rise either by the workers'
compensation benefit or by compensating wage premiums. If compensating wage premiums fully compensate for uninsured risk then changes in the benefit rates would have no effect on accident prevention or reporting; changes in the benefit rate would simply be offset by changes in the compensating wage premium that is paid for uninsured risk.

Employers also have no incentive to change or enforce their precautions against risk if compensating wage premiums fully compensate for uninsured risk. Higher benefit rates would reduce the uncompensated risk, and this would also reduce the compensating wage premium they have to pay for uninsured risks. Compensating wage premiums would adjust to make them indifferent between paying higher benefit rates or higher compensating wage premiums. Higher benefit rates would not induce them to increase their safety precautions because they would be paying lower compensating wages.

Empirical evidence on the existence of compensating wages for risk in Canada is scarce. The limited evidence, however, does suggest the existence of compensating wages for the risk of fatal injury. For the risk of non-fatal injury the evidence is more mixed, with small compensating wages found in two studies, and statistically insignificant ones found in one. These results tend to be in

---

4 Cousineau and Lacroix (1992) and Gunderson and Hyatt (1996).
line with those based on a larger number of U.S. studies \(^6\); that is, compensating wages do exist for risks of fatal injuries and more severe risks, but the evidence for less severe risk is weaker and more uncertain. Overall, the evidence suggests that there is some compensating wage premium paid for workplace risk, but that it is unlikely to fully compensate for the risk and hence to offset any incentive effects with respect to precautions or claims.

**Theoretically Expected Impact on Return to Work**

Basic economic theory predicts that, other things equal, an increase in an individual's non-labour market income or wealth would reduce their incentive to work since the additional wealth enables them to afford not to work. A reduction in their wage rate induces two types of behaviors that work in opposite directions with respect to return to work. On the one hand, a reduction in their wage reduces their income or wealth and hence they may have to work more to restore their lost wealth (the opposite of the income effect just discussed). On the other hand, a reduction in their wage reduces their monetary remuneration from work and hence provides less monetary incentive to work, making them more likely to substitute other activities (e.g., household activities, education, retirement, leisure) for work. These are respectively the income and substitution effect of a wage change, working in the opposite direction. Ultimately, it is an empirical question as to whether the income or substitution effect dominates the behavioral response of a group of workers.

---

As applied to workers' compensation benefits, this implies that disability rating schedule benefits (i.e., "meat chart" systems) would affect the incentive to return to work only insofar as the additional income would enable the recipient to afford not to have to return to work. Since the payment is not reduced if the person returns to work, then it does not discourage the return to work for that reason. In the parlance of economics, such lump-sum payments do not distort the decision margin between work and non-work activities; they have only pure income effects.

In contrast, wage loss systems effectively "claw-back" or reduce compensation as people return to work and restore their income. Individuals who are compensated for 90% of their lost net income, increase their earnings by only 10% if they return to work or increase their labour market work. In effect, such recipients of wage loss benefits face an implicit tax rate of 90% or higher.

In the parlance of economics, wage loss systems have both income and substitution effects that discourage return to work. The income effect reflects the fact that the compensation increases the income of recipients relative to what it would have been in the absence of the compensation, hence enabling the recipients to afford not to work or work as much. The substitution effect reflects the fact that the small or non-existent monetary return from working will induce the recipient to substitute into non-work related activities. In that vein, wage loss systems affect the return to work decision not only by increasing post transfer income (as do disability rating systems), but more importantly by reducing the monetary return to work. Wage loss systems distort the decision margin between work and non-work related activities by effectively eliminating the monetary incentive to work. The stronger adverse work incentive effects of wage loss systems occur because the compensation not only increase the recipients income (the income effect) but it also reduces the monetary returns to work (the substitution effect), while the disability rating systems only have the
income effect. This advantage of lump-sum payments under disability rating systems however must be traded off against the fact that they are not as good as wage loss systems in serving the insurance function of restoring lost income.

This greater adverse work incentive effects of wage loss systems is an inevitable by-product of the desire to have workers’ compensation serve its insurance function (i.e., to restore lost income) and yet to save on costs by reducing the compensation for those who return to work and restore their own lost income. If workers' compensation payments were not clawed back from those who restore their income, the costs of the system could be substantially greater than at present since recipients essentially would receive 90% of their former income as well as any additional labour market earnings they would earn. (These trade-offs are discussed in more detail under the program evaluation criteria). The policy challenge is to find the benefit system that serves the insurance function but that preserves the monetary incentive to return to work and restore lost income.

In such circumstances where the monetary incentives to return to work are reduced or eliminated, regulatory incentives invariably will be used to encourage -- indeed compel in many circumstances -- that return. Regulatory incentives, discussed previously, can come in such forms as return to work requirements, vocational rehabilitation requirements, reasonable accommodation requirements, and deeming provisions whereby workers who do not return to work are often deemed to be receiving the wage they could receive in a "suitable and available" job. In effect, the policy choice frequently is one of monetary incentives versus regulatory incentives, if one is to avoid reductions in the compensation programs themselves.
Summary of Theoretical Expected Impacts

Clearly, workers' compensation benefits can have potentially important effects on various outcomes, especially accident rates (real and reported) and return to work decisions. Table 2 summarizes these theoretically expected relationships. The top panel summarizes the expected effect of benefit rate increases on real accident rates and reporting of claims, as they are affected by both employees and employers. The bottom panel summarizes the expected effects on return to work decisions.

As indicated in the top panel, higher benefit rates are expected to reduce employee precautions because the costs are somewhat offset by the higher income replacement rates, but they are likely to increase employer precautions because the higher benefit payouts increase the costs of accidents, especially through experience rating. The net effect on real injuries is ambiguous depending upon whether the reduction in employee precautions or the increase in employer precautions dominates.

The effect of benefit increases on reported claims (as opposed to real injuries) is also ambiguous. Higher benefit rates may induce employees to report more claims from minor injuries, from off-the-job injuries, and from non-injury related problems, since these now may be compensated more generously through workers' compensation. These reporting effects are more likely to occur for hard-to-detect injuries, or if there is retrenchment in other sources of income support. Working in the other direction, employers are more likely to contest claims or reduce the reporting of accidents if the higher benefits would feed back to their costs, as would be the case for example with experience rating or other penalties. As with the true accident effect, the impact of higher benefits on the reporting and claims effect is theoretically indeterminate.
As indicated in the bottom panel, the impact of higher benefits on the monetary incentive to return to work depends in part on the type of compensation system in place. Under disability rating schedule systems the compensation would have a relatively modest impact on reducing the incentive to return to work; the reduced incentive occurs because the recipient has more income and therefore can better afford not to have to return. Under wage loss systems, however, the incentive to return can be reduced more substantially by the fact that the benefit is clawed back or eliminated if the worker returns. The effects on employers from the higher benefits are minimal under disability rating schemes because employer (workers’ compensation) costs are not substantially affected by any benefit reductions associated with the worker returning to work. Where experience rating policies are in place such that individual employer premiums more fully reflect claims costs, employer incentives to facilitate return to work is enhanced under a wage-loss pensions scheme. Overall, therefore, the incentive to return to work or restore lost income is not expected to be substantially affected by benefit increases under disability rating schedule systems, but they may be reduced under wage loss systems.

Some Ancillary Issues

The compensating wage differentials and work incentive analysis presented above can also provide insights into to debate on stacking versus integration of workers’ compensation benefits with benefits from other public and private programs. As described, compensating wage differentials for the risk of injuries will arise even in the presence of workers’ compensation if the workers’ compensation coverage is insufficient to cover the financial risk perceived by workers. If the risk the worker perceives is greater than that covered by workers’ compensation, the worker will seek, to the
fullest extent possible to cover that risk. The extent of the uninsured risk is in the eye of the worker — there is nothing in economic theory necessitating that total disability payments need be less than or equal to the worker’s earnings at the time of injury. Indeed, the uninsured risk as a result of a workplace injury may be greater to some workers than others if the worker has responsibilities or interests outside of work that the worker values, but are not reflected in their employment remuneration (e.g., care for an elderly parent which would have to financed by hiring the services of a nurse if the worker becomes injured). That disability benefits not exceed time-of-accident earnings is a product of insurance principles or other social notions of what is fair and equitable.

The term “compensating wage differential” is something of a misnomer in the sense that, while the additional remuneration received by the worker can come in the form of increased wages, it may be paid in the form of additional fringe benefits, including supplementary disability insurance. It would rarely be suggested that the worker was stacking benefits if they took their additional remuneration to compensate for the uninsured risk in the form of cash wages. Thus, when the worker takes part of the remuneration in supplementary disability benefits purchased on their behalf by the employer, or they use the additional money wages to purchase supplementary disability insurance on their own, it should be equally implausible to suggest that benefit stacking is inappropriate. The parties enter into these private arrangements to supplement short-comings in the workers’ compensation system to meet the risks they perceive. In short, their is little rationale from the perspective offered by the theory of compensating differentials for not permitting stacking of private benefits with workers’ compensation.

An additional concern is that the stacking of benefits from other sources with workers’ compensation is that it will reduce the incentive to return to work. As the preceding analysis
suggests, to the extent that disability benefits from any source are reduced by labour market earnings, the incentive to return to work is reduced because the implicit tax a return to work is high. Again, however, while additional benefits may increase the duration of absence from work for some workers, increases in benefit costs will also increase employer investments in safety, to the end of reducing the frequency of injuries, and increasing the incentive for employers to assist workers to return to work. Supplementary disability benefits provided through the employer are, in effect, fully experience rated. If workers’ compensation benefits are also experience rated, the employer faces incentives from both the private disability program and workers’ compensation to try to minimize these costs. Even if the worker purchases supplementary benefits on their own, the supplementary disability insurer and the workers’ compensation insurer both have incentives and responsibilities to monitor the progress of the claimant to ensure that periods of disability are not inappropriately extended.

Integration versus stacking of Canada Pension Plan benefits can be analyzed in a similar fashion. If a policy change was introduced to integrate CPP benefits with workers’ compensation benefits in a jurisdiction were CPP benefits were not previously integrated, then uninsured risks for some workers would increase. It would be expected that, over time, workers would demand compensating differentials in the form of higher wages or supplementary disability benefits to cover these now uncompensated risks. Thus, even though an employers workers’ compensation costs would fall, there is no guarantee that the total cost of labour would fall as a result of integrating CPP disability benefits with workers’ compensation benefits.
Similarly, to the extent that moving from a gross to net earnings basis for establishing workers’ compensation benefits results in reduced benefits to injured workers, and possibly increases in uninsured risk, it can be expected that other components of the worker’s total remuneration package would, over time, have to adjust to fill this new void.

EVIDENCE ON EFFECTS OF BENEFIT PAYMENTS

The previous section highlighted that workers' compensation benefits can have important effects on various outcomes, namely injuries (real and reported) and return to work. Often, however, the theoretical predictions were ambiguous, in part because employee and employer responses were expected to move in opposite directions. As such, it is important to look at the empirical evidence to determine both the direction and magnitude of any effect. Information on the magnitude of the effects is also necessary to inform policy choices given the trade-offs that are involved. For example, even if higher benefit rates do lead to a reduction in return to work or an increase in injury rates, that may be a "price" that society is willing to pay for more adequate coverage of injured workers, if those side-effects are small. This could be the case especially if the increase in injury rates were more from increased claims than from real injuries.

Empirical Methodology

The empirical methodology for estimating the impact of benefits on various outcomes generally involves the use of multiple regression where the relevant outcome is the dependent variable. Various outcomes have been utilized: injury rates, when using aggregate data; injury
probabilities, when using micro data based on individuals: injury duration; claim rates, probabilities and duration; and transitions from temporary to permanent disabilities. Sometimes, fatalities and medical injuries are used to distinguish between real and reported injuries on the grounds that reporting biases are unlikely to exist for fatalities, or for claims involving substantial medical expenses compared to those involving indemnity benefits.

The effect of workers’ compensation benefits is usually captured by including a measure of the generosity of benefits, such as the replacement rate (i.e., the ratio of benefits to wages). Other factors that can affect the outcome of interest are included as relevant control variables. When the dependent variable is a measure of duration (e.g., duration of injury or of time off work), hazard estimating procedures are sometimes employed. This essentially involves estimating the probability of exiting from the injured state or returning to work at any given day, conditional upon not having exited up until that point in time. This in turn can be reformulated to calculate expected duration and how they are affected by benefit rates.7

U.S. Results

Table 3 summarizes the results for a wide range of U.S. studies (and one study from the U.K. listed last). The earlier studies were generally based on aggregate data such as subsectors of manufacturing across different states and sometimes across different time periods. Variation in injury rates or duration were related to variation in benefit generosity across the states, industry subsectors

7 Using conventional regression, with the dependent variable entered as log duration, will give the same results as the more complicated duration hazard models if there are no time-varying covariates and if there is no right-censoring due to incomplete spells. Time varying covariates are those that can change over the duration spell. Censoring occurs if some of the observations reflect incomplete spells; that is, they are in the data set with a long duration but could still exit from the state (e.g., return to work).
and time periods. The later studies expanded on this in various ways: using micro data on individual workers; using hazard models to estimate duration; using fatalities and medical injuries to distinguish real and reporting effects; estimating transitions from temporary to permanent disabilities; and using unexpected changes in the benefit rates as "natural experiments" to estimate their effects. The last column gives the "bottom-line" results, with the phrase "increase" indicating that the benefit rate increased the outcome of interest such as injury or claim frequency or the duration of a claim.

Our interpretation of the generalizations that emerge from the U.S. studies is as follows:

• Higher benefit rates are generally associated with higher claims frequencies, an increased likelihood of conversion of temporary disability claims to permanent disability claims, as well as slower return to work.

• The evidence for duration of claims is more mixed although the more recent (and sophisticated studies) tend to find that higher benefit rates also increase such duration.

---

8 Summaries of this sort by necessity have to generalize especially when there are alternative specifications and the author(s) do not give strong statements of the preferred specification. As such, this should be regarded as a "broad-brush" picture. For specific details the original study should be consulted.

9 While most of the studies used outcomes that are generally construed as "undesirable" (e.g., injury rates) some used measures that are considered as "desirable" (e.g., labour force participation). For exposition purposes, these later outcome measures are transposed in the table so that an increase in the outcome is always considered as undesirable (e.g., labour force participation is transposed to out of the labour force or not having returned to work).
• There is some limited evidence that the reporting effects on increasing claims may be stronger than the real effects on increasing accidents, especially in the case of severe or fatal accidents. In fact, the real effects may well be to reduce severe accidents and fatalities\textsuperscript{10}.

• While the results could generally be interpreted as indicating that higher benefits have “negative” effects on most outcomes, the effects are often (but not always) quantitatively small. Most studies find elasticities of less than one, and often less than 0.5, although some find elasticities of greater than one\textsuperscript{11}. An elasticity of 0.4, which may represent the average, would imply that a 10\% increase in benefits, would result in a 4\% increase in claims. The substantial variation around that elasticity figure however (typically ranging from 0.1 to 1.6) suggests that one cannot say with confidence that the adverse impacts tend to be small.

• Even though the studies differ substantially in their methodology and the nature of their data, it is not obvious which of the studies are superior or inferior to others. Obviously, the more recent studies are likely to use the best methodology and the most relevant data, but their results tend to be similar to those of the earlier studies.

\textsuperscript{10} Moore and Viscusi (1990) find that higher benefit rates are associated with reduced fatalities, and Butler and Worrell (1991b) find that higher benefit rates are associated with reduced severe injuries that are associated with medical claims. These effects of higher benefit rates leading to reduced injuries, however, are not conventional. Worrell and Appel (1992) find that higher benefits increase claim duration but not for permanent or severe injuries. As indicated, many of the studies also find mixed results for the effect of benefits on the duration of claims, and duration would reflect the severity of the injury.

\textsuperscript{11} Typical rounded elasticities from the studies are 0.2 Chelius (1982), 0.4 Butler and Worrell (1983), 0.1 Johnson (1983), 0.5 Worrell and Appel (1982), 0.3 Butler (1983), 0.2 Ruser (1985), 0.2 Butler and Ward (1985), 0.3 Leigh (1985), 0.3 Chelius and Kavanaugh (1988), 0.0 Kneissner and Leeth (1989), 0.2 Krueger and Burton (1990), 0.7 Krueger (1990a), 1.6 Kruger (1990b), 1.1 Johnson and Ondrich (1990), 0.3 Butler and Worrel (1991b) [although 0.7 for claims and -0.4 for medical injuries], 0.3 Butler and Worrell (1991b), 0.3 Thomason (1993b), 1.1 for severe injuries and 0.1 for minor injuries Curington (1994), 0.4 Meyer, Viscusi and Durbin (1995), and 0.1 Fenn (1981). The average of those 20 elasticities is 0.4.
Canadian Results

Table 4 presents similar results from a smaller number of Canadian studies that examined the impact of workers' compensation benefits on various outcomes. Our interpretation of these results is as follows:

• In general, the results are similar to those of the U.S. studies; that is, higher benefit rates are generally associated with outcomes such as increased frequency and duration of claims, and increased likelihood of remaining out of the labour force.

• The Canadian results however are not uniform nor do they yield unambiguous conclusions. The relationship with claim duration, for example, is found to exist only in claims that are hard to diagnose\textsuperscript{12}. This could be expected, however, since it is difficult to monitor such cases. Some studies find that higher benefit rates reduce claims\textsuperscript{13} (contrary to U.S. evidence), although when that same data set is restricted to the more recent years the usual result of increased claims is supported\textsuperscript{14}. In some cases when the usual result of increased claims is found, it is marginally statistically insignificant\textsuperscript{15} or it applies to women but not to men\textsuperscript{16}.

\textsuperscript{12} Dionne and St-Michel (1991), Dionne, St-Michel and Vanesse (1995) and Fortin, Lanoie and Laporte (1996), all of which are based on Quebec data.  
\textsuperscript{13} Lanoie (1992) and Fortin and Lanoie (1992).  
\textsuperscript{14} Lanoie (1992b).  
\textsuperscript{15} Bruce and Atkins (1993).  
\textsuperscript{16} Johnson, Butler and Baldwin (1995).
• The ambiguity of the Canadian results is compounded by the fact that the benefit measures are often captured by imperfect proxies\textsuperscript{17} and the studies themselves do not discuss their own results in an unambiguous fashion\textsuperscript{18}.

• Given the ambiguity of the Canadian results perhaps a reasonable conclusion is that there is some tendency for higher benefits to be associated with increased claims and duration of such claims, although any impact is likely to be small. For example, a 10% increase in benefits is associated with a 4% increase in the likelihood of a claim\textsuperscript{19}, a 1% reduction in the likelihood of returning to work\textsuperscript{20}, and an 8% (i.e., 2.6 day) increase in the duration of claims\textsuperscript{21}.

\textsuperscript{17} Dionne and St-Michel (1991) and Dionne, St-Michel and Vanesse (1995) use a proxy for the extent of insurance coverage that simply captures whether the claimant has one of the following four characteristics: has a spouse in the labour force; has at least one dependent; worked part-time; and has an annual income less than the projected income used for compensation purposes. They argue that these individuals are likely to have insurance coverage that replaces more than 90\% of their net salary because of taxes and other features of the compensation program. Obviously this is a circuitous connection, and they realize that these variables may affect claim duration for other reasons.

\textsuperscript{18} Johnson, Butler and Baldwin (1995, p. 81) state "The benefit-wage ratio is not a significant influence on the duration of work absences among men. In the exponential model, the benefit-wage ratio significantly increases length of work absences amongst women ... The benefit-wage ratio was not a significant influence on the duration of work absence for men or women in the Weibul distribution." However, their tables themselves (p. 79) indicate that the benefit-wage ratio significantly increased the duration of work absences for women in the Weibel model, and it was positive but statistically insignificant in the exponential model. Further, they give no indication of whether the exponential or Weibel model is to be preferred.

\textsuperscript{19} Thomason and Pozzebon (1995).

\textsuperscript{20} Hyatt (1996).

\textsuperscript{21} Kralj (1995).
These results suggest that there is likely to be some trade-off between the insurance objective of improving the adequacy of income replacement, and other objectives such as reducing claims and their duration as well as encouraging the return to work. The trade-off does not appear to be extremely large, although this is not based on a large number of studies with unambiguous results.

The viability of reducing benefits to reduce the incidence and duration of claims and to encourage the return to work must also be judged against other mechanisms to meet these objectives. One study\(^{22}\) for example found that the return to work decision was much more responsive to an increase in wages — a 10% increase in wages would increase the likelihood of returning to work by approximately 10%, while a 10% reduction in the benefit amount would increase the probability of returning to work by only 1%.

EVALUATION ACCORDING TO ADEQUACY, EFFICIENCY AND EQUITY

The previous analysis provides information that can highlight the tradeoffs that are involved in choosing between disability rating schedules (meat charts) and wage loss systems for compensating for permanent partial disabilities, and for deciding upon the benefit replacement rate if wage loss systems are used. These tradeoffs can be evaluated according to the criteria of adequacy, efficiency and equity.

\(^{22}\) Hyatt (1996).
Adequacy, Efficiency and Equity Criteria

Adequacy refers to the extent to which the benefits adequately provide protection to injured workers. There are two dimensions to adequacy: covering all injured workers (the coverage dimension); and covering each of them to a sufficient degree (the degree dimension). A program could be inadequate because it fails to provide protection to all who merit protection, or because if fails to provide a sufficient degree of protection to those who are covered, or both.

Efficiency refers to the extent to which the benefit transfers interfere with the efficient allocation of resources so that the size of the pie is diminished. Inefficiencies occur when incentives are affected in a way that reduces the size of the economic pie to be distributed. In the area of workers’ compensation, the decision margins that can be affected in this way are the return to work decision, and the decision for both employers and employees to take adequate precautions against risk. If the return to work decision is adversely affected then output or the size of the economic pie is reduced. If adequate safety precautions are not taken, then that can be regarded as a direct reduction in the size of the pie since health and well-being are a component of that pie. In this area it is important to recognize that while the conventional emphasis tends to be on reducing disincentives that would otherwise discourage the return to work, it is also imperative to ensure that the incentives do not encourage a premature return to work that could subsequently jeopardize health and safety or other legitimate ends. It is also imperative to recognize that inefficiencies can result from excessive disputes and contesting of claims since such disputes involve the use of real resources that can be used in other areas. Disputes over how the pie is divided can eat up the pie itself. Lastly, in this area of efficiency it is important to emphasis efficiency is not a "be-all and end-all"; it is simply one of many objectives. It is entirely legitimate for society to chose a smaller economic pie that is more
equitably of fairly distributed; to make that choice, however, it is necessary to have information on the magnitude of any trade-off.

The criteria of equity or fairness refers to the distribution of the pie to be divided. There are two main dimensions to equity. Horizontal equity refers to the "equal treatment of equal" so that individuals who are in the same circumstances receive similar treatment. Vertical equity refers to the "unequal treatment of unequals" so that the disadvantaged are helped the most. Vertical equity, for example, would be facilitated by a "progressive" distribution of the benefits whereby benefits disproportionately went to the disadvantaged. There can be legitimate debate, however, over whether this criteria should be met through workers' compensation. Some may argue that it is an insurance program and not a redistributive social program. Others may argue that all programs should be judged somewhat according to this criteria, even if that is not their main purpose.

In applying the criteria to claims versus real accidents, it is important to make the distinction between real resource costs and transfer costs. Any effect of a compensation system on inducing more claims should not be weighted as heavily as the effect on real accidents. At worst, the induced claims are unmerited and even fraudulent, but this is largely a transfer cost from one group (payers) to another group (recipients). It is not a loss of real resources or human life or well-being, as is the case with the real accident effect. At best, the increased claims may be merited and even desirable. As discussed previously, the increased claims can come from various sources: reporting minor claims that otherwise would go unreported; reporting non-work claims as work-related claims; and reporting non-injuries as injuries. In many circumstances, however, these could be in the grey area of what constitutes a compensable claim, given the multiple causes of injuries and the complex interaction between work and non-work related causes.
While any increased transfer costs of the program may not be as serious as real resource costs (because with the transfer costs one party gains what the other party loses), it is also the case that the increased potential to transfer resources through claims can lead to costly litigation and other procedures to contest those claims or to check for fraudulent claims. This does represent a true loss of resources since those resources could be used for other purposes. This costly procedure in fact was one of the historical reasons for moving to a system of "no fault" workers' compensation where workers gave up their right to sue their employer in return for guaranteed and expedited compensation. In the United States, for example, where contingency fees are prominent in personal injury court cases, that fee is in the neighbourhood of 1/3 of the award. If the defendant spends a similar amount to defend against having to make the award, this suggests that over half of the amount transferred can be used up in the litigation process itself. This does suggest a "leaky bucket" in transferring from one party to the other, with the leaks going to law firms that do not exactly constitute the most disadvantaged members of society\(^23\).

While it is the case that transfer costs (e.g., from increased claims activity) are not as consequential as real resource costs (e.g., from increased real injuries), it is the case that transfers that result from fraudulent activity are more serious than transfers that result from legitimate mistakes or administrative difficulties in targeting payments to recipients. Of course, this requires information on the extent to which the claims are truly fraudulent, and such information is not readily available, albeit stories in both directions readily abound.

\(^{23}\) This may overstate the real resource cost of the process since particular cases can have precedence value that spills over to other cases that can be settled quickly without the costly litigation. Nevertheless, litigation and contesting claims is generally regarded as a costly process.
Disability Rating Schedules

Disability rating schedules or "meat chart" systems score poorly when evaluated according to the criteria of adequacy, at least from the insurance perspective where adequacy would be regarded as replacing lost income. The payment is related to lost earnings only in so far as the disability rating percent is multiplied by their pre-injury earnings to approximate their lost earnings capacity. But the rating itself does not differ depending upon the different effect the same injury may have on different people depending upon their circumstances or the nature of their work. It is designed to approximate the average reduction in earnings capacity, but there obviously may be considerable variation around that average depending upon how the injury affects the earnings capacity of a particular individual in their particular job.

With respect to meeting equity objectives, disability rating schedules score well with respect to horizontal equity (the equal treatment of equals) in that persons of the same earnings and with the same injury would receive the same compensation, subject of course to vagaries that could occur in assessing residual impairment. With respect to vertical equity (the unequal treatment of unequals) it would receive mixed grades. Since the disability rate of residual impairment is applied to pre-accident earnings, then that system is not "progressive" in that the benefits disproportionately do not go to persons of low earnings. For the system to be "progressive" a fixed benefit would have to apply, in which case it would be relatively larger for persons of low earnings. More importantly, the disability rating schedule treats individuals in a similar fashion even though a given injury may have a dramatically different effect on their ability to carry on working depending upon their occupation or the nature of their work. Equity in this vertical sense may require unequal treatment (differential compensation) of unequals (persons who are affected in a different fashion depending upon their occupation or the nature of their work).
With respect to meeting the efficiency criteria, disability rating schedules have the desirable feature that they do not distort decision margins with respect to the return to work decision. They affect that decision only in so far as they provide injured workers with additional income (relative to what they would have received without compensation) and that additional income may enable them to afford, for example, to delay their return to work or perhaps even not to return but to retire early. They are not discouraged from returning to work by the possibility that their workers' compensation payment would be reduced if they earned income. It is in part for this reason that a smaller payment can be involved, because the payment is not tied to any additional commitments such as a requirement to return to work.

With respect to affecting the decisions to take precautions against risk or to report injuries, the disability rating schedule would have the same effect as any compensation system that provided compensation benefits and hence reduced the costs of accidents and injuries. It is the magnitude of the benefits that matter for this decision and not whether they are tied to wage losses or are reduced if the recipient returns to work. Higher benefits may reduce the incentive of employees to take precautions against risk and to instigate more claims, but they may encourage employers, especially if experience rated, to take more precautions against risk and to contest claims. Our discussion of the empirical evidence suggested that higher benefits on net do tend to lead to more injuries and especially more claims. This consequence should be considered as an important trade-off in any decision to increase benefits.

As discussed previously, the increased claims are not as "serious" a problem as the increased injuries because the former are a transfer cost from one party to another, while the increased injuries are a real cost. The trade-off is made even more complicated by the possibility that increased benefits
lead to reduced injuries (as found in the two U.S. studies previously discussed with respect to fatalities and high medical cost injuries). Higher benefits leading to higher costs from greater claims activity may be a price worth paying (even if the claims are fraudulent) for reductions in severe accidents that lead to fatalities or injuries that involve major medical costs.

**Wage Loss Systems**

As discussed previously, wage loss systems should be regarded as a package or combination of compensation for wage loss and a set of administrative regulations designed to facilitate -- or perhaps compel -- the return to work. Those administrative regulations include vocational rehabilitation, return to work requirements, reasonable accommodation requirements, and deeming provisions. These regulatory dimensions subsequently will be assessed on their own according to the criteria of adequacy, equity and efficiency. At this stage only the wage loss component of that system will be assessed, but it is important to remember that it tends to be coupled with these regulatory components.

Wage loss systems score high when evaluated according to the criteria of adequacy, at least from the insurance perspective where adequacy would be regarded as replacing lost income. In fact, shifting from disability rating schedules to wage loss systems is generally rationalized on the basis of potentially being able to meet that insurance objective. The actual extent to which the adequacy aspect of meeting that objective is obtained of course depends upon the actual income replacement rate. The higher the replacement rate, the more likely the system will provide adequate benefits, but of course at the expense of conflicting with other objectives (discussed subsequently). Furthermore, for persons above the maximum, the replacement rates decline markedly, and the system therefore does not serve the insurance principle of replacing lost income.
With respect to the criteria of equity or fairness, wage loss systems can do well in terms of horizontal equity or the "equal treatment of equals" since benefits are designed to be in proportion to losses (e.g., replace 90% of lost income). In theory, persons who have the same income and who experience the same injury (or an injury that leads to the same income loss) will receive the same compensation.

Wage loss systems with a constant replacement rate would not meet the vertical equity objective of "unequal treatment of unequals" since low income persons would have the same proportion of their income replaced as would high income persons. However, a substantial degree of progressivity (i.e., disproportionately helping low income persons) can be attained through the minimums and maximums. The minimums lead to low wage persons who are at or below the minimum receiving 100% of their former earnings. The maximums lead to persons who are above the maximum receiving less than their nominal replacement rate. Their actual replacement rate can be substantially less since the maximum is a fixed maximum -- the replacement rate itself does not simply decline slowly as earnings increase. If the maximum, for example, were $500 per week, and the replacement rate were 90%, then a person just at that maximum would have a benefit payment of $450 per week. If the person formerly earned $1000 per week (around $50,000 per year) then they would receive the maximum of $500 or an actual replacement rate of 50%. Clearly, such a system has a substantial degree of progressivity.

While such progressivity facilitates vertical equity, it conflicts with insurance principles of replacing lost income. Persons who insure 90% of the value of their house or car do not receive 50% of that loss simply because they have above-average earnings. As well, if progressivity is to be attained there is a question of whether it should be done so on the basis of individual earnings (as is
the case with workers' compensation) or family earnings since the later tends to be a better measure of economic need. Also there is the question of whether progressivity should be attained through the workers' compensation system (which was designed more as an insurance system and not a redistribution system) or through other mechanisms, notably the income tax system where the "progressive" tax rates are explicitly designed to achieve a degree of progressivity, and where dependents and other dimension of need are explicitly considered. This raises the question of the rationale for exempting workers' compensation payments from income taxes when the income tax system is designed in part to tax according to need and to redistribute income. If the maximums under workers' compensation are designed to provide a degree of progressivity, would that not be better achieved by allowing workers' compensation payments to be taxable? Higher average payments could be made to offset the loss of the former tax advantage if that were deemed appropriate. Certainly, the extremely progressive implicit tax rates that are involved for persons above the maximums in workers' compensation are much higher than those in the income tax system, especially because the maximums apply for persons around the average earnings level. This suggests, at a minimum, that more consideration should be given to the rationale for the maximums, and if that rationale pertains to progressivity, whether that could be better achieved through the income tax system.

With respect to the criteria of allocative efficiency, wage loss systems would likely receive a poor score at least with respect to the dimension of encouraging the return to work. Obviously the higher the replacement rate the less the monetary incentive to return to work. As discussed previously, replacement rates in the neighbourhood of 90% provide almost no monetary incentive to return to work. Recipients increase their earnings by only 10% by returning to work or augmenting
their earnings; in effect, they are subject to 90% implicit tax rates or clawbacks. In such circumstances, the monetary incentives are replaced by administrative incentives through regulations such as requirements pertaining to vocational rehabilitation, return to work, reasonable accommodations and deeming provisions. As such, the efficiency aspects of wage loss systems must be evaluated as part of the package that also includes these as components. Those components are evaluated separately next.

With respect to the dimension of encouraging precautions against real risk and discouraging "excessive" (e.g., fraudulent) claim activity, wage loss systems again would receive a relatively poor score at least to the extent to which they had generous replacement rates. As discussed previously under the disability rating schedule system, the incentives in this area depend upon the magnitude of the benefits and not whether they are tied to wage losses or are reduced if the recipient returns to work. Higher benefits may reduce the incentive of employees to take precautions against risk and to instigate more claims, but they may encourage employers, especially if experience rated, to take more precautions against risk and to contest claims. Our discussion of the empirical evidence suggested that higher benefits on net do tend to lead to more injuries and especially more claims. This consequence, however, is tempered by the fact that the increased claims are a transfer cost from one party to another. Furthermore, the evidence on increased injuries is not unambiguous since a few studies find higher benefits to be associated with fewer fatalities and medical injuries.

Administrative Regulations

As indicated previously, administrative regulations should also be regarded as an integral component of workers' compensation benefits. In particular, wage loss compensation systems should
be regarded as being coupled with a set of administrative regulations designed in part to facilitate -- indeed compel -- restoring lost income and returning to work when feasible. This is so because when there are few monetary incentives to return to work, then administrative regulations will be used as a substitute. The administrative regulations include vocational rehabilitation, return to work requirements, reasonable accommodation requirements, and deeming provisions.

With respect to the criteria of adequacy of benefits, the administrative regulations can be regarded as designed to offset the adverse incentive effects of high benefit rates. In that sense, they could be thought of as facilitating adequate benefits. Yet they are also intended in part to encourage and compel the return to work, and in that sense they may be regarded as providing inadequate benefits if there is undue or inappropriate pressure to return to work so as to reduce benefits. Certainly, the deeming provisions are often regarded by potential recipients as thinly veiled excuses for cutting benefits.

With respect to the criteria of horizontal equity (the equal treatment of equals), the administrative requirements likely score poorly. Administrative requirements always run the risk of somewhat arbitrary treatment since they tend to be applied or enforced by different individuals (and because application of the rules can be adjusted over time to meet other objectives of the organization). Even the same individual may not apply the rules and regulations in a consistent fashion over time and across different cases. Biases and differences in perceptions are always possible. The difficulties are highlighted by the key phrases that are often involved with interpreting regulations: "reasonable accommodation" up to the point of "undue hardship"; "deeming" to be able to earn the wage of "suitable and available employment"; returning the injured worker to the pre-accident job or a "comparable" job or a "suitable" job. All of these involve difficult judgement calls

37
and hence raise the possibility that otherwise similar workers will receive different treatment depending upon how those calls are made.

With respect to the criteria of vertical equity (the unequal treatment of unequals) that is usually associated with disproportionately helping the disadvantaged, the administrative requirements also likely score poorly. The disadvantaged are also likely to be similarly disadvantaged in dealing with bureaucratic, administrative regulations, although advocacy groups can help them in that regard. It is easy to imagine, for example, that in determining what is a "suitable" job for re-employment purposes or for deeming purposes, that the principle that "any job will do" will often be applied to disadvantaged workers.

With respect to the criteria of allocative efficiency, regulatory requirements also likely score poorly. This is so especially when they use up real resources that are therefore not available for other purposes. In essence, many of the regulatory requirements exist to prevent abuses in the system -- abuses that are more likely in wage loss systems with high replacement rates. Such abuses can include fraudulent claims activities and a reluctance to return to work or to restore lost income on the part of persons whose claim depends upon the continuation of lost income. Regulatory requirements associated with vocational rehabilitation, return to work obligations, reasonable accommodation requirements and deeming provisions are designed in part to substitute for the absence of monetary incentives. But these regulatory requirements are expensive. They reduce efficiency by reducing the resources otherwise available to society. The pie is shrunk by the regulatory requirements that are established in part to deal with its distribution in this area. To the extent that these regulatory requirements foster adversarial relations and lead to abuses in their administration, then this could also be regarded as reducing well-being.
Summary of Criteria Applied to Different Systems

Table 5 provides a summary picture of the evaluation criteria of adequacy, equity and efficiency applied to the disability rating schedule system compared to the wage loss system, when the wage loss system is also associated with extensive administrative, regulatory requirements that must also be evaluated according to the same criteria.

The disability rating schedule system scores poorly in terms of adequacy of the benefits according to the insurance principle of replacing lost income. Wage loss system scores well in this regard since that is their main purpose, although for persons above the maximum (i.e., above the average wage) actual replacement rates drop considerably and wage loss systems score poorly in terms of replacing lost income.

Both systems score well in terms of horizontal equity since they tend to treat similar workers in a similar fashion. The administrative regulatory requirements that tend to be associated with wage loss systems, however, score poorly in this area since there is always a degree of arbitrariness in their application and interpretation.

Disability rating systems score poorly in terms of vertical equity since they do not disproportionately help the disadvantaged, while the wage loss systems have a high degree of progressivity through the minimums and especially the maximums. It is questionable, however, as to whether such a high degree of progressivity is intended, and whether an insurance system like workers’ compensation should be called upon to serve such a redistributive function, or if that would be best done through the tax system. The regulatory requirements that are associated with wage loss
systems, however, score poorly in terms of vertical equity in that the disadvantaged are likely to be disadvantaged in dealing with regulatory and administrative bureaucracy.

Disability rating systems score highly in efficiency since they minimize disincentives to return to work and restore lost income. Wage loss systems are extremely poor in this regard because there is little monetary incentive to return to work. Regulatory requirements are designed in part to substitute for the lack of monetary incentives, but they tend to use up real resources that could be used for other purposes.

ELEMENTS OF AN “OPTIMAL” BENEFIT STRUCTURE

The previous summary highlights that no system unambiguously dominates in the sense of meeting all or even most of the criteria of adequacy, equity and efficiency. Tradeoffs abound.

To a large degree, the choice between systems based on disability rating schedules versus wage losses depends upon one's view of the importance of the lack of monetary incentives to return to work and restore lost income in the wage loss systems, and the need to replace those lost incentives with administrative, regulatory ones that have other negative effects. As well, it depends upon one's view of the appropriateness of the substantial degree of progressivity that is obtained through the minimums and especially the maximums in the wage loss system. While progressivity is desirable, that degree of progressivity is surprisingly high (much higher than that of the regular income tax system) and it conflicts with the insurance objective of adequately replacing lost income for substantial numbers who are beyond the average wage. It is also questionable as to whether an insurance scheme should be called upon to serve that redistributive function or whether it would be best done through the tax system.
The previous analysis also highlighted that neither theory nor empirical evidence provided unambiguous answers so as to guide choices on the trade-offs. As indicated previously in Table 2, when the (usually offsetting) actions of both employers and employees were considered, higher benefits could increase or decrease real and reported injuries. Furthermore, disability rating schedules could have an ambiguous effect on the return to work decision, while wage loss systems would unambiguously reduce the incentive to return to work, setting up the need for complementary regulatory requirements to substitute for those reduced monetary incentives.

The empirical evidence provided some help in ascertaining the magnitude of the tradeoffs, although the results are sometimes conflicting (albeit no more so than in most areas of empirical research). In general, higher benefit rates are associated with outcomes such as increased injuries and especially claims and duration of claims, although any impact is likely to be small. The results suggest that there is some trade-off between the insurance objective of improving the adequacy of income replacement, and other objectives such as reducing claims and their duration as well as encouraging the return to work. Some limited Canadian evidence suggested that the return to work decision was much more responsive to an increase in wages, highlighting the importance of that dimension.

While this degree of ambiguity in both the theoretical predictions and the empirical evidence does not provide a clear blue-print for reform in this area, it does provide some insights that may not have been obvious without the theoretical and empirical exercise. Those insights include:

- Higher benefits reduce precautions from employees but increase them for employers, hence they have an ambiguous effect on injuries.
• Similarly, higher benefits can induce employees to report more injuries to enhance claims, but they induce employers to contest the claims and hence have an ambiguous reporting effect on claims activities.

• Disability rating systems have a relatively mild effect on reducing the monetary incentive to return to work, while wage loss systems substantially reduce that monetary incentive.

• The reduced monetary incentive in wage loss systems lead to more extensive regulatory requirements in such areas as vocational rehabilitation, return to work obligations, reasonable accommodation requirements, and deeming provisions. Eliminating monetary incentives inevitably leads to more use of regulatory requirements.

• For that reason, the wage loss systems are tied to more extensive regulatory requirements. The choice between disability rating schedules and wage loss systems are really ones where the wage loss systems should be regarded as coupled with the more extensive regulatory requirements. This detracts from their otherwise attractive feature of being able to provide adequate insurance protection against income loss, at least for persons of below average wages who would not be subject to the maximums.

• The minimums and maximums under wage loss systems impart a substantial degree of progressivity so that income replacement can easily be 90 to 100% of lost income for low-wage individuals, but less than 50% of lost income for persons of moderately higher income. This also means that persons above the average wage and hence above most maximums, are not well "insured" against income loss. There is also some question as to whether such an insurance system should serve a redistributive function or if that would best be left up to the tax system. To that end, more consideration could be given to not exempting workers' compensation from income taxes.
While most of the evidence suggests that higher benefit rates are associated with outcomes such as increased real and especially reported injuries, increased duration of claims, and slower return to work, the evidence is by no means unambiguous. A few studies found reductions in real injuries, especially serious ones like fatalities and injuries associated with medical expenses.

While lower benefit rates in wage loss systems may be merited to reduce these outcomes, there are caveats and qualifications that temper and may even reverse that conclusion: the evidence is by no means conclusive; higher claims rates (which unambiguously occur) involve transfer costs whereby what one party looses another party gains; and such higher transfer costs may be a price worth paying for increased adequacy of coverage and especially if they were associated with a reduction in severe injuries as found in a small number of studies.

These theoretical and empirical insights suggest that the search for a unique, optimal benefit rate in wage loss systems is likely to be at best an elusive search. Some insights into that search may be gleaned from a study that dealt specifically with that issue (and to our knowledge the only study that has attempted to do so). Viscusi and Evans (1990) use information on wage premiums for risky work to infer what individuals would be willing to pay to purchase actuarial fair insurance to replace lost income from an injury. They calculate the optimal income replacement rate to be 0.85 of gross income based on the fact that individuals have revealed that they would not purchase full insurance because their ability to enjoy their income is reduced if they are injured\(^ {24} \). They further calculate the optimal earnings replacement rate would be reduced to 0.68 because about $0.20 of every dollar goes to the costs of administering the system. They note that this optimal replacement rate is close to the effective replacement rate of about 0.64 that existed in the U.S. at that time.

\(^ {24} \) Whether this occurs is an empirical proposition since income could also matter more to persons if they become disabled so as to better cope with the disability.
Hyatt (1996) followed a different approach but came to surprisingly similar conclusions. He used information from permanent disability recipients in Ontario prior to 1990 (i.e., when Ontario paid permanent disability awards based on functional impairment) to estimate the impact of expected post-injury labour market wages and permanent disability benefits on the likelihood that these workers were employed at the time they reached maximum medical recovery. He found, as expected, that the likelihood of employment increased with increases in post-injury wages, and decreased with increases in the permanent disability benefit. Hyatt then asked the hypothetical question, how much would labour market earnings need to be increased by to offset the disincentive created by giving the injured worker an additional dollar of disability benefits. The answer was between $1.10 and $1.24, depending on which factors were held constant in the statistical analysis. In other words, Hyatt’s sample of workers in Ontario were willing to “pay” (out of foregone labour market earnings) $1.10 to $1.24 in labour market earnings for an additional dollar of workers’ compensation benefits. In a sense, following their injury, workers value a dollar of workers’ compensation benefits more than they value a dollar of labour market earnings.

This can be restated to provide an estimate of an appropriate replacement rate. In this instance, the relevant question is how much of their workers’ compensation benefit would a worker be willing to give up in return for an additional dollar of labour market earnings. The answer is the between 80 cents and 91 cents (i.e., the inverse of $1.24 and $1.10). This suggests that a defensible replacement rate, based on this approach which analyzes the implicit valuation of benefits by workers, is between 80 percent and 91 percent of net earnings.

While these studies should not be taken as a precise derivation of the true optimal replacement rate, they do highlight that the optimal replacement rate is less than full income
replacement because individuals would not voluntarily buy full replacement insurance because their ability to enjoy the replacement income is (unfortunately) reduced by their injury, and there are substantial administrative costs associated with providing such insurance.

Clearly, an optimal replacement rate remains an elusive search. The issue is complicated by the fact that the optimal rate is influenced by a wide range of considerations: the desire to use it to both cover lost income and to serve a redistributive rationale; the ability to enjoy income in the injured state; the administrative costs of the system; and its impact on real and reported injuries and the return to work decisions.

This paper highlighted the complexities of the many trade-offs that were involved. Its intent was also to provide some theoretical and empirical insights that may not be "common" knowledge in this complex area, and to show how the different parties may adjust to changes in benefits. Hopefully, this will narrow the range of debate and provide some common ground for determining the appropriate benefit structure for workers' compensation.
REFERENCES AND WORKS CITED


<table>
<thead>
<tr>
<th>Jurisdiction</th>
<th>% of Earnings</th>
<th>Maximum $</th>
<th>Minimum $</th>
</tr>
</thead>
<tbody>
<tr>
<td>British Columbia</td>
<td>75% of gross</td>
<td>803</td>
<td>277</td>
</tr>
<tr>
<td>Alberta</td>
<td>90% of net</td>
<td>554</td>
<td>232</td>
</tr>
<tr>
<td>Saskatchewan</td>
<td>90% of net</td>
<td>620</td>
<td>258</td>
</tr>
<tr>
<td>Manitoba</td>
<td>90% of net²⁶</td>
<td>561</td>
<td>134</td>
</tr>
<tr>
<td>Ontario</td>
<td>90% of net</td>
<td>691</td>
<td>303</td>
</tr>
<tr>
<td>Quebec</td>
<td>90% of net</td>
<td>594</td>
<td>245</td>
</tr>
<tr>
<td>New Brunswick</td>
<td>80% of net²⁷</td>
<td>474²⁸</td>
<td>none</td>
</tr>
<tr>
<td>Nova Scotia</td>
<td>75% of net²⁹</td>
<td>414³⁰</td>
<td>none</td>
</tr>
<tr>
<td>Prince Edward Island</td>
<td>80% of net³¹</td>
<td>391</td>
<td>none</td>
</tr>
<tr>
<td>Newfoundland</td>
<td>75% of net³²</td>
<td>458³³</td>
<td>200</td>
</tr>
<tr>
<td>Yukon</td>
<td>75% of gross</td>
<td>782</td>
<td>none</td>
</tr>
<tr>
<td>NWT</td>
<td>90% of net</td>
<td>649</td>
<td>310</td>
</tr>
</tbody>
</table>


²⁵ B.C., Alberta, Saskatchewan and the NWT provide for 100% of net earnings if less, and Manitoba provides for 90% of net earnings if less.
²⁶ 80% of net after 24 months of cumulative benefits.
²⁷ 80% for first 39 weeks, thereafter 85%.
²⁸ 474 for married and 451 for single based on 80% replacement; 504 for married and 479 for single based on 85% replacement after 39 weeks.
²⁹ 75% for first 26 weeks, 85% thereafter.
³⁰ 414 for 75% replacement, 469 for 85% replacement.
³¹ 80% for first 39 weeks, 85% thereafter.
³² 75% for first 39 weeks, 80% thereafter.
³³ $458 for 75% replacement, $489 for 80% replacement after 39 weeks.
TABLE 2 -- THEORETICALLY EXPECTED IMPACT OF BENEFIT RATES ON INJURIES (REAL AND REPORTED) AND ON RETURN TO WORK

<table>
<thead>
<tr>
<th></th>
<th>Employees</th>
<th>Employers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Impact on Injuries</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real (moral hazard effect)</td>
<td>Reduce precautions</td>
<td>Increase precautions</td>
<td>Ambiguous</td>
</tr>
<tr>
<td>Reported (claims effect)</td>
<td>Increase claims</td>
<td>Contest claims</td>
<td>Ambiguous</td>
</tr>
<tr>
<td><strong>Impact on Return to Work</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability rating schedule system</td>
<td>Reduce slightly (income effect)</td>
<td>Increase slightly</td>
<td>Ambiguous</td>
</tr>
<tr>
<td>Wage loss system</td>
<td>Reduce (both income and substitution effects)</td>
<td>Increase slightly</td>
<td>Reduce</td>
</tr>
</tbody>
</table>

Note: The analyses assume that firms are not fully experience rated and compensating wages do not fully compensate for risk.
### TABLE 3 -- U.S. STUDIES OF THE IMPACT OF WORKERS' COMPENSATION BENEFITS ON VARIOUS OUTCOMES

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Data</th>
<th>Methodology</th>
<th>Outcome</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chelius (1977)</td>
<td>18 states, 17 ind., 1967</td>
<td>Regression</td>
<td>Injury rate</td>
<td>Increase</td>
</tr>
<tr>
<td>Chelius (1982)</td>
<td>36 states, mfg., 1972-75</td>
<td>Regression</td>
<td>Injury rate</td>
<td>Increase Decrease</td>
</tr>
<tr>
<td>Chelius (1983)</td>
<td>28 states, mfg., 1972-78</td>
<td>Regression</td>
<td>Injury rate Duration Total lost days</td>
<td>Increase No effect Increase</td>
</tr>
<tr>
<td>Chelius &amp; Smith (1983)</td>
<td>37 states, 15 industries 1979</td>
<td>Regression</td>
<td>Injury rate</td>
<td>Increase</td>
</tr>
<tr>
<td>Johnson (1983)</td>
<td>432 NY workers, 1975</td>
<td>Probit</td>
<td>Out labor force</td>
<td>Increase</td>
</tr>
<tr>
<td>Butler (1983)</td>
<td>15 S. Carolina industries, 1940-69</td>
<td>Regression</td>
<td>Claim frequency</td>
<td>Usually increase, but often insig.</td>
</tr>
<tr>
<td>Butler &amp; Worrell (1983)</td>
<td>35 states, 1972-78</td>
<td>Regression</td>
<td>Claim frequency</td>
<td>Usually increase</td>
</tr>
<tr>
<td>Leigh (1985)</td>
<td>11,889 individuals, PSID, 1977-79</td>
<td>Regression</td>
<td>Claim probability</td>
<td>Increase</td>
</tr>
<tr>
<td>Ruser (1985)</td>
<td>41 states, 25 industries, 1972-79</td>
<td>Regression</td>
<td>Injury rate</td>
<td>Increase</td>
</tr>
<tr>
<td>Butler &amp; Worrell (1985)</td>
<td>Illinois men, back injuries</td>
<td>Duration hazards</td>
<td>Claim duration</td>
<td>Increase</td>
</tr>
<tr>
<td>Moore &amp; Viscusi (1990)</td>
<td>1,173 individuals, 1982 PSID</td>
<td>Regression</td>
<td>Fatalities</td>
<td>Decrease, especially if experience rated</td>
</tr>
<tr>
<td>Krueger (1990a)</td>
<td>CPS individuals 1983-85</td>
<td>Regression</td>
<td>Probability of claim</td>
<td>Increase</td>
</tr>
<tr>
<td>Reference</td>
<td>Location, Time Period</td>
<td>Methodology</td>
<td>Outcome</td>
<td>Findings</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Krueger (1990b)</td>
<td>Minnesota injuries</td>
<td>Log duration, diff. in diff.</td>
<td>Claim duration</td>
<td>Increase, small if experience rated</td>
</tr>
<tr>
<td>Johnson &amp; Ondrich (1990)</td>
<td>1040 individuals, 3 states, 1970</td>
<td>Duration hazards</td>
<td>Claim duration</td>
<td>Increase</td>
</tr>
<tr>
<td>Krueger &amp; Burton (1990)</td>
<td>States, 1972-82</td>
<td>Regression</td>
<td>Claims</td>
<td>Mixed or no effect</td>
</tr>
<tr>
<td>Ruser (1991)</td>
<td>2800 establishments, 1979-84</td>
<td>Regression &amp; poison reg.</td>
<td>Injury rate</td>
<td>Increase, smaller if exp. rated</td>
</tr>
<tr>
<td>Butler &amp; Worrell (1991a)</td>
<td>12 states</td>
<td>Duration regression</td>
<td>Claim duration</td>
<td>Increase</td>
</tr>
<tr>
<td>Butler &amp; Worrell (1991b)</td>
<td>33 states, 1954-81</td>
<td>Regression</td>
<td>Claims cost, medical cost</td>
<td>Increase claims, reduce injuries</td>
</tr>
<tr>
<td>Thomason (1993b)</td>
<td>Individuals, 5 states, 1979-81</td>
<td>Tobit &amp; twin probability</td>
<td>Transition temp. to perm.</td>
<td>Increase</td>
</tr>
<tr>
<td>Carrington (1994)</td>
<td>N.Y. state, 1963-83</td>
<td>Log duration &amp; difference in difference</td>
<td>Claim duration</td>
<td>Increase especially for severe injuries</td>
</tr>
<tr>
<td>Fenn (1981)</td>
<td>U.K. survey of 2,142 injured workers</td>
<td>Illness duration</td>
<td></td>
<td>Increase</td>
</tr>
</tbody>
</table>
TABLE 4 -- CANADIAN STUDIES OF THE IMPACT OF WORKERS' COMPENSATION BENEFITS ON VARIOUS OUTCOMES

<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Data</th>
<th>Methodology</th>
<th>Outcome</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dionne &amp; St-Michel (1991)</td>
<td>Injured workers, Quebec 1978-82</td>
<td>Log duration regression</td>
<td>Duration of claim</td>
<td>Increase for hard to diagnose cases</td>
</tr>
<tr>
<td>Lanoie (1992a)</td>
<td>Quebec industries 1974-87</td>
<td>Regression</td>
<td>Claim rates</td>
<td>Reduce</td>
</tr>
<tr>
<td>Fortin &amp; Lanoie (1992)</td>
<td>Quebec industries 1974-87</td>
<td>Regression</td>
<td>Claim rates</td>
<td>Usually reduce</td>
</tr>
<tr>
<td>Lanoie (1992b)</td>
<td>Quebec industries 1983-87</td>
<td>Regression</td>
<td>Claim rates</td>
<td>Increase</td>
</tr>
<tr>
<td>Dionne, St-Michel &amp; Vanesse (1995)</td>
<td>Quebec WCB survey and admin. data 1987</td>
<td>Log duration regression</td>
<td>Duration of claim</td>
<td>Increase for hard to diagnose cases</td>
</tr>
<tr>
<td>Fortin, Lanoie &amp; Laporte (1996)</td>
<td>Quebec construction workers 1976-86</td>
<td>Duration hazards</td>
<td>Duration of claim</td>
<td>Increase for hard to diagnose cases</td>
</tr>
<tr>
<td>Criteria</td>
<td>Disability Rating Schedule (&quot;Meat Chart&quot;)</td>
<td>Wage Loss System</td>
<td>Administrative Regulations</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Adequacy</td>
<td>Poor in replacing lost income</td>
<td>Good in replacing lost income but poor for persons above maximums</td>
<td>Designed in part to deal with potential abuses of high benefit rates</td>
<td></td>
</tr>
<tr>
<td>Horizontal Equity (Equal Treatment of Equals)</td>
<td>Good since similar workers given similar treatment</td>
<td>Good since similar workers receiving similar wage loss replacement</td>
<td>Poor since can be applied arbitrarily</td>
<td></td>
</tr>
<tr>
<td>Vertical Equity (Unequal Treatment of Unequals)</td>
<td>Poor, does not disproportionately help disadvantaged or those whose disability leads to greater wage loss</td>
<td>Good in that minimums and especially maximums help low wage workers more, although questionable as to whether best means for such redistribution</td>
<td>Poor in that disadvantaged likely to be disadvantaged in dealing with regulatory requirements</td>
<td></td>
</tr>
<tr>
<td>Efficiency</td>
<td>Good, minimizes disincentive to return to work or restore lost income and few resources used in administering regulations. Some adverse effect in reducing safety precautions</td>
<td>Poor, little incentive to return to work or restore lost income; poor in that reduces safety precautions and encourages inappropriate claims activity</td>
<td>Poor since uses up real resources in administering and implementing regulations</td>
<td></td>
</tr>
</tbody>
</table>