Occupational Safety and Health: a Problem of Double or Single Moral Hazard

Paul Lanoie

ABSTRACT

Following Cooper and Ross (1985), this article contributes to the literature by developing a principal-agent approach to the effect of imperfect information on safety in the workplace. Firms and workers can influence the probability of workplace accidents by taking precautions. To the extent that a worker's precaution cannot be observed by a firm, and vice versa, contract designing is a problem characterized by moral hazard. Double and single moral hazard lead to nonoptimal levels of precaution from firms and workers, but not necessarily to an underprovision of precaution. Furthermore, it is shown that regulatory policies aimed at promoting safety in the workplace (by inducing increases in the levels of precaution) may not be welfare-improving.

Some authors, such as Thaler and Rosen (1976), allege that the presence of wage differentials for risky jobs implies that public regulation of occupational safety and health (OSH) is unnecessary. They argue that risk premiums in dangerous occupations are enough to compensate a worker in the case of an accident and to induce the socially optimal effort needed to reduce hazards.[1] This argument against regulation of safety in the workplace holds in a world of perfect markets and complete information. However, it breaks down with incomplete information or imperfect markets, and the literature on accident prevention[2] shows that nonoptimal situations can occur. For instance, there are externalities; Manga et al. (1981) argue that risk in the workplace leads to indirect costs that should be taken into account in the firm's decision concerning safety expenditures. An obvious example is the indirect health cost that a child incurs when a pregnant woman is at risk.

---

*Paul Lanoie is Assistant Professor, Institute d'Économie Appliquée, École des Hautes Études Commerciales, Montréal, Canada.

Financial support from the Social Sciences and Humanities Research Council of Canada and from the Fonds F.C.A.R. is gratefully acknowledged. The author is indebted to Richard Arnott, Charles Beach, Lorne Carmichael, George Dionne, Carmen Pedro, Dan Usher and two anonymous referees for helpful comments.

[1] There is plenty of empirical evidence that firms pay a risk premium to workers in risky occupations; see, for instance, Thaler and Rosen (1976) and Viscusi (1978). However, the empirical evidence is contestable, see Olson (1981). Lanoie (1990) examines theoretically the circumstances under which the existence of risk premia is more likely.

[2] For a review of this literature, see Manga et al. (1981) and Viscusi (1980, 1983).

Problems of imperfect information in the "market of workplace accidents" have also drawn special attention. In particular, Oi (1974), Diamond (1977) and Rea (1981) postulate that workers are not well informed about safety levels prevailing in different firms. Since market outcomes are not efficient in this case, the effect of various policies such as mandatory insurance (workers' compensation) and safety regulation should be analyzed. Diamond and Oi argue that mandatory insurance is justified because it raises the expected utility of risk averse workers, while the enforcement of safety standards is required in order to improve the level of safety in the workplace, which is suboptimal when workers underestimate risk. In a different model, Rea (1981) shows that, when workers underestimate risk, mandatory insurance and safety regulation may have the perverse effect of reducing welfare and the level of safety. In particular, wages may be reduced with mandatory insurance and, as a result, the level of safety may fall because workers will attempt to substitute wages for safer jobs.[3] Carmichael (1986) disagrees with the "uncomfortably ad hoc" assumption that workers underestimate risk. He integrates imperfect information by exploring the role of a firm's reputation in repeated games. His model suggests that it takes time for workers to learn about changes in safety levels in a firm which leads, generally, to an underprovision of safety. In contrast with Rea, Carmichael is able to make unambiguous statements about the welfare-improving nature of government intervention related to OSH. Among other results, he shows that an increase in the level of compensation benefits received in accident cases leads to an unambiguous improvement in welfare.

This article contributes to the literature by working out a new approach to the effect that imperfect information has on safety in the workplace. In particular, the problems of double and single moral hazard are examined in the "market of workplace accidents". Moral hazard is a generic phenomenon defined as the effect of insurance on the choice of self-protection activities by the insured when the insurer cannot observe or enforce these activities. Single-sided moral hazard typically leads to an underprovision of self-protection activities by the insured, at least when one concentrates on the substitution effect. The expression "double moral hazard" is borrowed from Cooper and Ross (1985) who examine the important characteristics of warranties (including their provision of incomplete insurance and the relationship between product quality and coverage) in a related analysis where buyers and sellers take actions that affect a product's performance.[4] Similarly, a workplace accident depends on the precaution levels of both the firm and the worker. Therefore, the provision of insurance by the firm could create a

[3] There is no such ambiguity in Diamond's results because he assumes that safety is constant when he analyzes the insurance decision, and that insurance benefits are constant when he analyzes the safety decision.

[4] For pioneering work on moral hazard, see Pauly (1974), Holstrom (1979), and Arnott and Stiglitz (1982). For another treatment of double moral hazard related to product guarantees, see Kambhu (1982) and Emons (1988). Mann and Wissink (1988) examine double moral hazard in the context of money-back contracts. For a treatment of bilateral accidents (car accidents) with moral hazard, see Boyer and Dionne (1987) and Gravelle (1986).

double moral hazard problem if it affects the (nonenforceable) precaution levels of the two parties to the contract. In fact, a rise in the level of insurance benefits gives opposite incentives to both parties; it decreases the cost of an accident to the worker (inducing less precaution), while it increases the cost of an accident to the firm (inducing more precaution). As a result, double moral hazard does not necessarily lead to an underprovision of precaution by both parties; whether it does depend not only on the substitutability or the complementarity of the precaution levels of the two parties, as in Cooper and Ross (1985), but also on the chosen level of insurance.

Moral hazard problems are modelled through the interaction within a principal-agent framework, of a risk neutral firm and a risk averse worker in which both parties, by taking precautions, can influence the probability of workplace accidents. The first step in the analysis is to construct a cooperative first-best solution with full information that can serve as a benchmark. The nature of the moral hazard problem can be understood by comparing the market equilibrium levels of compensation benefit (insurance) and accident-prevention effort (provided by firm and worker) with their first-best levels. The market equilibrium (a noncooperative solution) is modelled as a two-stage game, and two game equilibrium concepts are chosen to illustrate two noncooperative solutions, the Nash and the Stackelberg equilibria. Both equilibrium concepts seem relevant within a labor market context. In the Nash equilibrium, the precaution levels of both the firm and the worker are nonenforceable in a contract (double moral hazard). This type of equilibrium corresponds to a situation in which, for instance, a firm cannot monitor an employee taking drugs or alcohol, and a worker cannot observe elements such as the level of toxic substances in the air or the level of noise. It is an environment in which illnesses (such as deafness) are more likely to be contracted because of the unobservability of a firm's precaution. In the Stackelberg equilibrium, the firm's precaution is enforceable (single moral hazard) since, under certain circumstances (e.g., a firm installs safety guards on a machine), it is likely to be more "visible" than a worker's risk-preventing activities. In such an environment, injuries instead of illnesses are more likely to occur.

Both noncooperative equilibria are analyzed and compared. They are both nonoptimal, and it is shown how government intervention might reduce the welfare loss that arises when workers and firms fail to cooperate. The model indicates, however, that regulatory policies aimed at promoting safety in the workplace (by inducing increases in the levels of precaution) might not be welfare-improving.

This article extends the work of Cooper and Ross to the context of the labor market. Specifically, it differs from their contribution in three important respects. First, both agents in their model are risk neutral so that their warranties are designed ignoring risk-sharing considerations. Second, they do not consider the Stackelberg equilibrium which is relevant in the context of the labor market. Third, they neither make a welfare analysis nor discuss the policy implications of their results. The current analysis shows that Cooper and Ross' results do not generalize to the case where only one of the two economic agents is risk neutral or to the case where the informational set defines a Stackelberg equilibrium. Furthermore, the relationship between first-best and noncooperative levels of precaution is explored in more detail, which allows more precision about the circumstances leading to the under or overprovision of precaution by economic agents. Compared to other studies in the OSH literature, the risk of accident can be influenced by both the firm and the worker, which is a more general modelling assumption than in Carmichael's work where risk depends only on the firm's resources devoted to safety. As a result, his unambiguous policy recommendations do not extend to the model presented here. Although Rea considers the possibility that both a firm's and a worker's precaution could affect the risk of an accident, he does not examine the case in which a firm's risk-preventing behavior is not observable by workers. Moreover, the case of safety regulation with workers' precaution being unobservable is analyzed, which is neglected in Rea's article, and different circumstances are found under which safety regulation might be welfare-decreasing.

The rest of the article is organized as follows: The next section introduces the basic model and notation, and briefly presents the full-information solution. Then, noncooperative or asymmetric information cases are examined and followed by an exploration of how government intervention can mitigate the moral hazard problem. Finally, concluding remarks are given.

The Model

Consider a representative risk-neutral firm in a competitive industry where there is a risk of workplace accidents. The firm, which has no capital, hires a fixed number of identical risk-averse employees each working a fixed number of hours during the period and making one unit of product.

It is important to specify the timing of the problem. The events in this static model occur within one period, but their timing is not the same. The worker receives a wage, w, per period and chooses an accident-prevention effort, e, per period over a continuum of non-negative effort levels. This effort can be interpreted in terms of time spent in risk-preventing activities or the unpleasantness of these activities (e.g., wearing safety glasses). The firm selects a non-negative safety expenditure per worker of q per period. Examples of this kind of expenditure would be safety equipment distributed to workers or investment in safer technology. These variables should be thought of as occurring continuously throughout the period (discounting is ignored). There may or may not be an accident. For simplicity, it is assumed that there is only one type of accident of fixed severity. The accident, if it occurs, takes place on the last day of the period and does not, therefore, affect a worker's output. The firm is then considered strictly liable[5] and has to pay the worker an ex

[5] For reasons not specifically modelled here, such as adverse selection, it is assumed that workers cannot insure themselves in a private market. Since the firm is risk-neutral, one would expect it to be the natural source of insurance.

post compensation benefit, α, on that day[6]. Our timing of the accident differs from that of other authors (e.g., Rea (1981), Carmichael (1986)), who implicitly assume that an accident can happen only on the first day of the period.[7] The decision to place the accident on the last day is no more extreme than theirs, and it is intended as a representation, within a one-period model, of the fact that receipt of wages is continuous while accidents are rare and discrete.

The probability of an accident, P, is a decreasing function of e and q:

$$P = P(q,e), \qquad P_i < 0,\ P_{ii} > 0 \text{ for } i = e,q, \text{ and } 0 \leq P(e,q) \leq 1. \tag{1}$$

For analytical convenience, it is also assumed that $P_{eq} = 0$, signifying that an increase in the level of one safety input, e or q, has no effect on the marginal productivity of the other safety input. As mentioned by Rea (1981), some expenditures by the firm may inform employees (signs, for example) and act as complements to workers' effort ($P_{eq} < 0$). However, other precautions taken by the firm, such as safety devices on a machine, are likely to be perceived as substitutes for workers' effort ($P_{eq} > 0$).[8] In this text, the main results are based on the assumption that $P_{eq} = 0$ although discussion is provided on how the results are affected by P_{eq} being positive or negative.

The worker's expected utility, EU, depends on consumption if there is no accident (consumption is assumed to equal income), compensation in the event of an accident, effort, and the probability of an accident. With appropriate simplifying assumptions, EU becomes:

$$EU = P(q,e)U^a + (1 - P(q,e))U^n - e. \tag{2}$$

The utility from consumption is event-dependent.[9] U^a is the worker's utility function if there is an accident, and U^n is the utility function if there is none. Since w and α are received at different times in the period, it is not unreasonable to postulate that U^a is additively separable in w and α.[10] Therefore, the function U^a is assumed to be the sum of the worker's utility during the period (or before the realization of the uncertainty), U(w), which depends only on w, and his or her utility on the last day, $Z^a(\alpha)$, which depends only on α; i.e., $U^a(w,\alpha) = U(w) + Z^a(\alpha)$. Similarly, $U^n(w) = U(w) + Z^n$, where Z^n is the utility obtained on the last day if there is no accident and it is

[6] It would be plausible to assume that α is a function of the wage since α is usually thought of as compensating the worker for a loss of income. Instead, a general case is considered, as in Rea (1981) and Carmichael (1986), where α would compensate the worker for a loss of utility.

[7] Krueger (1988) adopts similar accident timing in his two-period model where an accident can happen only at the beginning of the second period, and the disabled workers produce no output in the second period.

[8] Carmichael (1984) also assumes that $P_{eq} = 0$.

[9] For a complete discussion of moral hazard and event-dependent utility functions, see Dionne (1982).

[10] It is conventional to assume that utility of income is separable through time (e.g., Svensson and Razin (1983)).

assumed to be a constant[11]. Utility functions are assumed to be continuous, twice differentiable and strictly concave in consumption.[12] Furthermore, only the situations in which a worker is not better off because of an accident are considered; i.e., it is postulated that $Z^n > Z^a$ in equilibrium (which implies that $U^n > U^a$). Finally, since the effort is undertaken before it is known whether or not the accident occurs, the disutility of effort is assumed, as in Arnott and Stiglitz (1982), to be event-independent. For brevity, P, U^a and U^n are often used without their arguments in the text.

The firm maximizes expected profit per worker:

$$E\pi = f - w - q - P\alpha \tag{3}$$

where f, the price of the good, is assumed to be constant.

The first step in this investigation of moral hazard problems is to construct the full information solution. In this solution, firm and worker set all the elements of the contract (w,α,e,q) "cooperatively", and the contract is fully enforceable. To obtain the full information solution, one maximizes $E\pi$ subject to $EU = \overline{U}$, where $\overline{U}$ may be thought of as the equilibrium utility of a worker in the entire labor market;[13] i.e., the utility that could be obtained by working in another industry. The cooperative solution (w^*, α^*, e^*, q^*) satisfies the first-order conditions (F.O.C.):[14]

$$w: \quad -1 + \lambda U_w = 0 \tag{4}$$

$$\alpha: \quad -P + \lambda P U^a_\alpha = 0; \qquad \lambda = 1/U_w = 1/U^a_\alpha > 0 \tag{5}$$

$$e: \quad P_e(-\alpha)/\lambda + P_e(U^a - U^n) = 1 \tag{6}$$

$$q: \quad P_q(-\alpha) + \lambda P_q(U^a - U^n) = 1 \tag{7}$$

$$\lambda: \quad PU^a + (1-P)U^n - e - \overline{U} = 0 \tag{8}$$

From (4) and (5), it is clear that α^* and w^* are set so that the marginal utilities of income in the accident and non-accident event are equal ($U^a_\alpha = U^n_W$), which implies full monetary insurance.[15] Furthermore, as in Cooper and Ross (1985),

---

[11] Since this model examines moral hazard problems, the analysis is concentrated on insurance and levels of care (α,e,q). The timing of the accident and separability in U^a and U^n are therefore intended to neutralize, for tractability, the impact of w in the model. Cooper and Ross (1985) also neutralize the influence of one of their endogenous variables by linearizing their problem.

[12] That is, $U^j(\bullet)$ for $j = a, n$ is a function from R_+ into R with $U^j_y(\bullet) > 0$ and $U^j_{yy}(\bullet) < 0$ for $y = w, \alpha$.

[13] In their cooperative solution, Cooper and Ross (1985) maximize the sum of the seller's expected profits and the buyer's expected utility. Another possibility would be to maximize expected utility subject to a firm break-even constraint but, as the reader can verify, doing so would not alter the qualitative nature of the results.

[14] The reader can verify that the second-order conditions (S.O.C.) are satisfied.

[15] This is a standard result also obtained by Rea (1981) and Carmichael (1986). However, it is not found in Cooper and Ross (1985) because of the linearization of their problem.

e^* and q^* are set so that the sum of the marginal benefits to both parties equal the marginal costs to each. This rule is analogous to the efficiency condition in public goods theory.

The Noncooperative Solution

Moral hazard problems can arise when there is noncooperative interaction between firm and worker in which e and/or q are not necessarily enforceable in a contract. Indeed, firms may have difficulties monitoring a worker's risk-related behavior, such as running on a wet floor, while employees may be unable to control certain elements such as the constant quality of the ambient air. Noncooperative interaction between firm and worker is modelled as a two-stage game played by both parties. As indicated earlier, two game equilibrium concepts are considered, the Nash and the Stackelberg equilibria. In the Nash (N) case, both q and e are chosen noncooperatively by the firm and the employee respectively in the second stage of the game and therefore, there is double moral hazard. However, w and α are still determined cooperatively, and in an enforceable manner, in the first stage. This modelling technique is somewhat realistic since, in actual labor contracts, firms and workers bargain mainly over wages and other types of compensation rather than over their respective safety inputs. The Stackelberg case will be treated subsequently.

Nash Equilibrium

As in dynamic programming, one first examines the second stage of the game for arbitrary levels of w and α. When safety inputs q and e are chosen noncooperatively, the moral hazard problem[16] can be analyzed by looking for reaction function equilibria. For a given w and α, and knowing how P varies with e for a given q, the worker chooses e to maximize (2). The solution to this problem, $e = e(w,\alpha)$, is derived from the first-order condition (to avoid excessive notation, the superscript N (for Nash) is dropped when there is no ambiguity):

$$e\colon P_e(U^a - U^n) - 1 = 0. \tag{9}$$

Similarly, the firm maximizes its expected profits per worker with respect to q, given w and α, and knowing how P varies with q for a given e. The solution $q = q(w,\alpha)$ satisfies:

$$q\colon P_q(-\alpha) - 1 = 0 \tag{10}$$

[16] As explained in Arnott and Stiglitz (1982), the moral hazard problem may give rise to non-convexities in indifference curves and in the set of feasible insurance policies, which might undermine the existence of a competitive equilibrium. These difficulties are ignored throughout the article. As a consequence, the noncooperative equilibria (Nash and Stackelberg) are assumed to exist, to be unique and stable. For a discussion on multiple equilibria in the context of double moral hazard, see Cooper and Ross [(1985), p. 112].

Totally differentiating (9) and (10), one can deduce respectively the slopes of the worker's ($e = f(q)$) and the firm's ($q = h(e)$) reaction functions in e,q space:

$$\left(\frac{de}{dq}\right)^f = -\frac{P_{eq}}{P_{ee}} = 0 \tag{11}$$

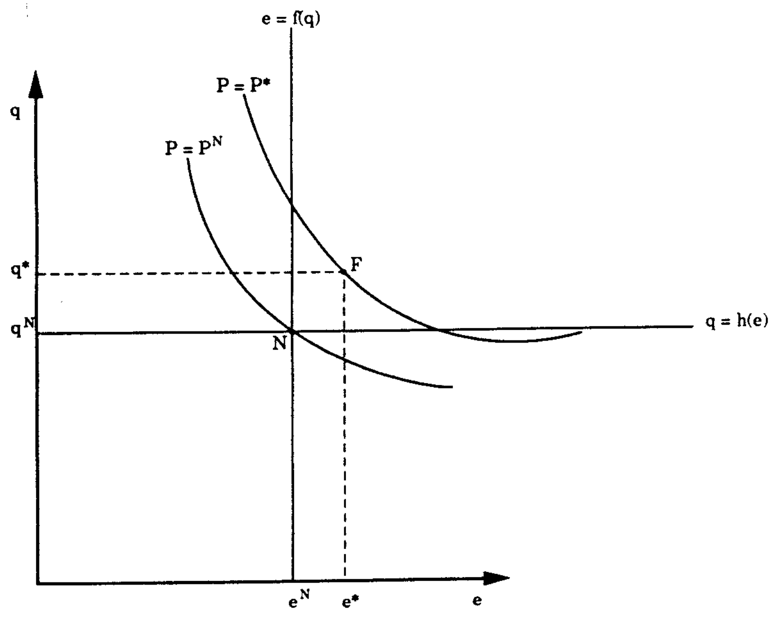
$$\left(\frac{dq}{de}\right)^h = -\frac{P_{qe}}{P_{qq}} = 0 \tag{12}$$

In e,q space, the worker's reaction function has an infinite slope and the firm's a zero slope (see Figure 1 below). This means that when one party increases its safety-enhancing action, the other's action remains unchanged. This result depends crucially on the assumption that safety inputs are independent ($P_{eq} = 0$). Both reaction functions would have a negative slope if e and q were perceived as substitutes and a positive slope if they were perceived as complements.

The case where $e^N < e^*$ and $q^N < q^*$, for arbitrary levels of α and w, is illustrated in Figure 1. Worker and firm reaction functions intersect at N (the Nash equilibrium). The first-best solution F is depicted at e^*, q^*, which lies on the isoprobability locus $P = P^*$. The negative slope of this locus is deduced from total differentiation of $P = P(e,q)$. Similarly, the Nash solution N is depicted at e^N, q^N, which lies on the isoprobability locus $P = P^N$.

Figure 1:

The Nash Equilibrium



Isoprobability loci further away from the origin (involving higher levels of precaution) correspond to lower probabilities of accident.

It is instructive to compare carefully the Nash equilibrium with the first-best solution. Equation (9) shows that, when the worker selects e noncooperatively, his or her only concern is the effect of a change in e on his or her own welfare. There is no consideration for the effect of a change in e on the firm's expected compensation costs (compare (9) with (6)). Likewise, when picking q noncooperatively, the firm is only concerned with the payment of α in the case of an accident (compare (10) with (7)). Therefore, in contrast to the cooperative setting, the worker chooses e and the employer chooses q to equate marginal cost to his or her own marginal benefit. This suggests that both e^N and q^N are lower than their corresponding levels in the first-best solution as in a standard externality problem. Although appealing, this intuitive conclusion is not completely correct. It can be shown, however, that e^* and q^* must differ from e and q in the Nash equilibrium. More specifically:

Proposition 1:

1.A. If $q^* = q^N$, then $e^* > e^N$
1.B. If $e^* = e^N$, then $q^* > q^N$
1.C. It is not possible that $e^N > e^*$ and $q^N > q^*$ simultaneously.
Proof: See Appendix I.

The reader can verify that this proposition also holds when $P_{eq} > 0$. When $P_{eq} < 0$, the proposition holds if $\alpha^N = \alpha^*$; otherwise, the complementarity or the synergy between the two safety inputs ($P_{eq} < 0$) may understandably make both agents "too" prudent in the Nash equilibrium. Proposition 1 confirms the suboptimality of the Nash solution since the first-best levels of care cannot be attained simultaneously. It shows that double moral hazard does not necessarily lead to the underprovision of precaution from both parties. This outcome depends not only on the substitutability or the complementarity of the precaution levels of the two parties as suggested by Cooper and Ross (1985), but also on the chosen level of insurance compared to the first-best level, which was not pointed out in the existing literature. Furthermore, non-linearities induced by a worker's risk aversion are important since, with both agents being risk neutral, Cooper and Ross show that e^N and q^N are necessarily smaller than e^* and q^* respectively when $P_{eq} = 0$.

To confirm that the outcome depends on the chosen level of insurance, one can solve for the first stage of the Nash game in which w and α are chosen cooperatively given $e(w,\alpha)$ and $q(w,\alpha)$. The solution, presented in Appendix II, shows that, in the Nash case, full monetary insurance would only occur fortuitously. Indeed, there is no obvious reason why it should be the case, or why α^* should be greater or smaller than α^N. Knowing this, it is useful to reexamine Proposition 1. From this Proposition, it is clear that two other cases have to be ruled out to confirm the validity of the intuitive proposition, $e^* > e^N$ and $q^* > q^N$. These cases are: (i) $q^N > q^*$ and $e^N < e^*$, and (ii) $q^N < q^*$ and $e^N > e^*$. They cannot be ruled out because of double moral hazard;

that is, because a variation in α from its first-best level has a different impact on the precaution levels of the two parties. Two situations may be distinguished. First, suppose that $\alpha^N > \alpha^*$, then a greater level of insurance is provided in the Nash equilibrium than in the first-best solution. Better insurance coverage reduces the opportunity cost of an accident for the worker, and increases the opportunity cost of an accident for the firm which pays the insurance, inducing the worker to be less careful ($de/d\alpha > 0$ from (9)) and the firm to be more careful ($dq/d\alpha > 0$ from (10)). Consequently, when $\alpha^N > \alpha^*$, it is conceivable that the firm chooses $q^N > q^*$, while the worker chooses $e^N < e^*$ (as in case (i)). Conversely, when $\alpha^N < \alpha^*$, it is conceivable that the worker selects $e^N > e^*$, while $q^N < q^*$ (as in case (ii)). It follows that the intuitive case in which $e^N < e^*$ and $q^N < q^*$ occurs when the Nash equilibrium α is the same as the first-best α, so that there are no contradictory effects of a change in α on the two safety inputs. In fact, it can be shown that:

Proposition 2:

2.A: If $q^N > q^*$ and $e^N < e^*$, then $\alpha^N > \alpha^*$.
2.B: If $q^N < q^*$ and $e^N > e^*$, then $\alpha^N < \alpha^*$.
2.C: $q^N < q^*$ and $e^N < e^*$ when $\alpha^N = \alpha^*$.
Proof: See Appendix I.[17]

Stackelberg Equilibrium

As discussed in the introduction, the Stackelberg equilibrium, where one party takes into account the other's reaction function, is also an appropriate game equilibrium concept to illustrate the interaction between firm and worker in an environment where injuries instead of illnesses are more likely to occur. This interaction is now modelled as a two-stage game where the safety expenditure, q, the compensation benefit, α, and the wage, w, are determined by the firm in the first stage of the game given $e = e(q,w,\alpha)$ (i.e., given the worker's reaction function), and taking the workers' interests into account by providing them with their reservation level of utility $\overline{U}$. Workers then pick their effort, e, in the second stage taking q, α and w as given. In other words, the modelling assumptions now imply that q, α and w are enforceable in a contract between firm and workers, while e is left out of it and therefore, there is single moral hazard. These assumptions seem realistic because, in many circumstances, the safety expenditure, q, precedes employment (e.g., the design of a machine), and may be considered as "visible" capital. A worker's effort, however, is at best imperfectly observable (although firms can certainly influence it). Hence, it is natural to think of the firm as committing itself by being the first actor and to think of the workers as reacting (for instance,

[17] The reader can verify that, as in Proposition 1, Proposition 2 does not necessarily hold when $P_{eq} > 0$.

workers could react by using the safety equipment put at their disposal, or they could become more careful if the firm is negligent, etc.).[18]

In the Stackelberg setting (S), the worker's second-stage noncooperative maximization problem is still characterized by equation (9). In the first stage, the firm selects w, α and q to maximize expected profits subject to EU $= \bar{U}$ and to the worker's reaction function, e(q,w,α). The solution to this first stage problem is provided in Appendix II. It shows that, as in the Nash equilibrium, a full insurance outcome would only occur fortuitously.[19]

Concerning the levels of precaution, since the worker's choice of e is still characterized by equation (9), one would expect that e^s should be lower than e^* (compare (6) with (9)). However, there is a countervailing effect; as a Stackelberg leader, the firm has a natural incentive to reduce its (enforceable) safety expenditure, and induce the worker to provide more safety input in the P function. This last result, in turn, can be counterbalanced by the fact that, when deciding upon its safety expenditures, q, the firm now takes into account the worker's marginal benefit of one more unit of q (by providing the worker with his or her reservation level of expected utility), which was not the case in the Nash equilibrium. Altogether, these conflicting effects may interact in many different ways so that the relationship between the precaution levels chosen in the Stackelberg equilibrium and the first-best levels is more complex. However, when one assumes, as in Proposition 2, that the Stackelberg equilibrium α is equal to the first-best level, it can be shown that:

Proposition 3: When $\alpha^* = \alpha^S$, then $e^S < e^*$ and $q^S > q^*$.
Proof: See Appendix I

and

Proposition 4: When $\alpha^* = \alpha^S$ and $P_{eq} > 0$:
if $q^S = q^*$, then $e^S < e^*$
if $e^S = e^*$, then $q^S < q^*$
it is not possible that $e^S > e^*$ and $q^S > q^*$ simultaneously.
Proof: See Appendix I.

Proposition 3 indicates that, when $\alpha^S = \alpha^*$ and $P_{eq} = 0$, the precaution level chosen by the worker in the Stackelberg equilibrium is necessarily smaller than the first-best level, while the firm is "over-prudent". This result is new in the literature and differs from the Nash equilibrium where both precaution levels are shown to be lower than first-best levels with $\alpha^* = \alpha^N$ and $P_{eq} = 0$.

[18] As pointed out by a referee, the critical aspect of this noncooperative solution is that the firm's risk-preventing action is now observable by the worker, and can be made part of the contract. This results in a single moral hazard problem. Other sequences in the decisions of the parties are conceivable.

[19] Compare equations (II.5) and (II.6) in Appendix II with (4) and (5) respectively. Furthermore, by rearranging (II.6): $U^a_\alpha = 1/\lambda^S + (P_e\alpha(de/d\alpha))/P\lambda^S$, it is clear that the greater in absolute value the depressing effect of α on e, the lower is α. Therefore, if α has a strong influence on e at the margin, the firm reduces α to stimulate the worker's effort.

An explanation for this result is that, when an increase in q does not affect the marginal productivity of e ($P_{eq} = 0$), and when the firm cannot influence the worker's precaution through a change in α ($\alpha^S = \alpha^*$), the firm is induced to choose a higher level of safety expenditure than the first-best level knowing that the worker will choose a lower level of precaution than at the optimum (because of moral hazard).

Proposition 4 is exactly analogous to Proposition 1, except that, because of the complex relationship between the Stackelberg and the first-best levels of precaution (discussed above), one has to make an additional assumption ($\alpha^* = \alpha^S$) to prove the Proposition. Lanoie (1989b) shows that neither Proposition 3 nor 4 necessarily hold if $P_{eq} < 0$. Indeed, as in the Nash equilibrium, the complementarity between the two types of precaution may induce both parties to be "over-prudent".

The Stackelberg (S) and first-best (F) equilibria are illustrated in Fig. 2 under the assumption that $P_{eq} = 0$ and $\alpha^S = \alpha^*$, so that $e^S < e^*$ and $q^S > q^*$ (see Proposition 3). The employee's and the firm's reaction functions would have a negative slope if e and q were substitutes (as will be shown in Figure 3), and a positive slope if they were complements. Furthermore, the shape of the isoprofit locus is obtained from the total differentiation of equation (3), whereas the shape of the indifference curve is obtained from the total differentiation of equation (2).[20]

Neither the Stackelberg nor the Nash equilibrium is Pareto optimal. Representing different working environments, they are both second-best solutions and there is no a priori basis in welfare for preferring one to the other. Since both (N and S) are essentially of the same nature, the rest of the analysis focuses on the Stackelberg equilibrium, which is neglected by Cooper and Ross (1985).

Government Intervention

Can a government intervene to push the economy from the Stackelberg equilibrium, with inappropriate levels of care and compensation benefit, to the Pareto optimum? The answer to this question depends on the information that the government has at its disposal. Technological constraints make the case of a perfectly informed government unlikely. Therefore, a more relevant question is: can a government with no more information than economic agents induce a Pareto improvement over the noncooperative solution? In the context of the present model, one natural way to tackle this question is to ascertain whether actual policies adopted by governments, which are likely to have no more information than agents, are welfare improving.

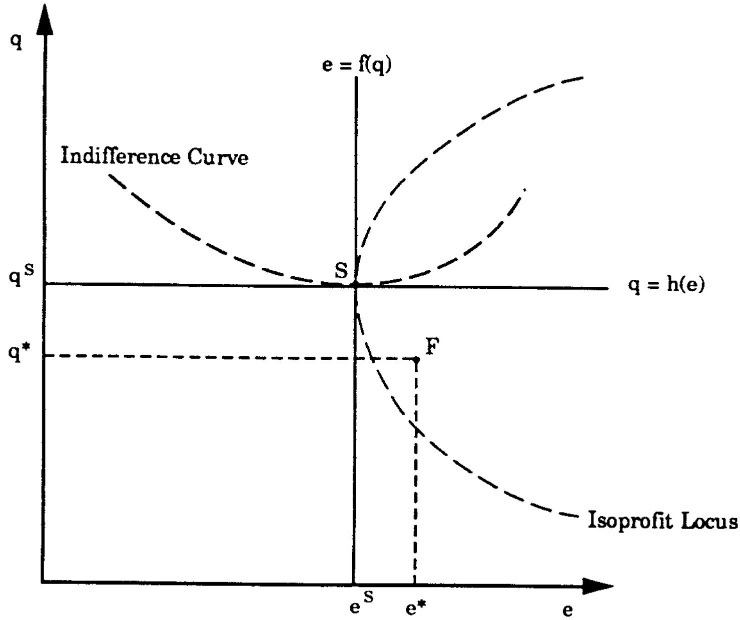
In most North American jurisdictions, compensation α is determined by the government and safety standards are imposed to promote safety in the workplace. Although some of these standards address worker conduct (e.g.,

---

[20] It should be noted that e^S and q^S are not necessarily equal to e^N and q^N respectively in Figure 1.

Figure 2

The Stackelberg Equilibrium



do not speed while driving a forklift truck), most are imposed on firms and stipulate either equipment specifications (e.g., width of railings) or equipment performance (e.g., ladder rungs must be able to support 350 lb.). Different measures are also implemented for the enforcement of standards (inspections, fines and prosecutions).[21]

Two avenues are envisaged for policy analysis. First, the welfare implications of safety-enhancing measures are analysed and second, the effect of a change in compensation benefits. To clarify and simplify the presentation of the arguments, simple graphs are used and the following assumptions are made: (i) α is fixed by the government, (ii) w is not affected by safety policies,[22] (iii) the sum of expected profits and utility is maximized in the first stage of the Stackelberg game, and (iv) the government can act at no cost.

Concerning policies intended to promote safety, situations where safety inputs are independent and substitute ($P_{eq} = 0$ and $P_{eq} > 0$) are examined.[23] With $P_{eq} = 0$, Proposition 3 shows that $q^S > q^*$ and $e^S < e^*$ (see Fig. 2). The implications of this result are that there is no need for policies that induce firms to devote more resources to safety (in fact, firms should be induced to take less precaution), and that OSH authorities should concentrate attention

[21] For more institutional details, see Lanoie (1989b).

[22] Lanoie (1989b) provides empirical support for this assumption.

[23] The case in which $P_{eq} < 0$ is more difficult to analyze since it was not possible, in Section 3, to derive precise relationships between the first-best levels of precaution and the Stackelberg levels.

on worker incentives. The latter prescription is probably feasible; in Canada, for instance, OSH boards provide workers with specific safety incentives through the adoption of such measures as the right to refuse a dangerous task or the compulsory creation of joint worksite safety committees in certain high risk industries.

When $P_{eq} > 0$, it can be shown that, under certain circumstances, it is advantageous to raise e and q, but the increase must remain "within certain bounds" to improve welfare. This point can be made by referring to Figure 3, which depicts the worker's and the firm's reaction function (in the e,q space) by loci $e = f(q)$ and $q = h(e)$ respectively. IC and IP represent respectively the worker's indifference curves and the firm's isoprofit loci with expected utility and profits increasing when moving away from the origin to the northeast.[24] This implies that the point F is a welfare-improving point compared to an initial Stackelberg equilibrium (S).

Consequently, when policy instruments such as those used in North America (safety standards, right of refusal and others) are implemented to make the firm and the worker more cautious (thus shifting reaction functions to the right),[25] any incremental move from S to the northeast is sufficient to induce a welfare improvement. However, this is only true as long as q and e are increased in appropriate proportions. One of the implications of this statement is that OSH authorities should avoid excessive standards that would correspond to a move from S to B. For instance, Mendeloff (1979) suggests that American health standards might not be welfare-improving. When the OSHA (Occupational Safety and Health Administration) sets new health standards, it must decide whether exposure to toxic substances or noise has to be limited by the use of engineering controls, or whether firms can rely on personal protective equipment such as earmuffs or respirators. In most cases, engineering controls are adopted (which correspond to a large increase in safety expenditure, q), whereas "cheap" personal protective equipment could have fulfilled the same task. Clearly, such a policy is not necessarily welfare-improving since a higher level of safety is reached, though at a very high cost. Rea (1981) finds that safety regulation might also be welfare-decreasing if firms react to the regulation by reducing wages.

However, one should be aware that it is not necessary that both e and q be increased for a Pareto improvement. Strictly speaking, there is nothing in the model that precludes F′, in Figure 3, from being a welfare-improving point over S, and F′ corresponds to a situation where $e^* < e^S$ and $q^* > q^S$. For instance, this case could occur when an unduly negligent firm induces its workers to be over-prudent. Although such a situation may be common, it is unlikely that any central authority would ever perform a calculation showing

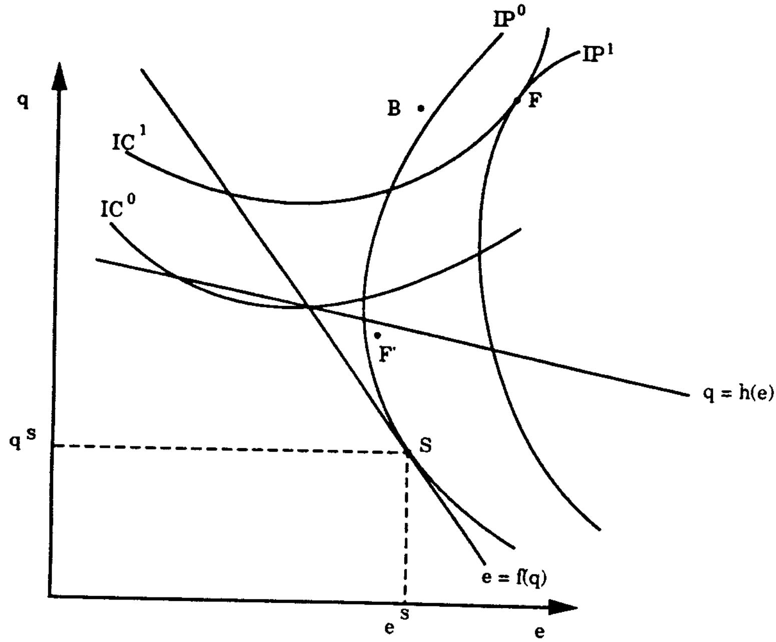
---

[24] It is clear from (3) that, ceteris paribus, $\partial E\pi/\partial e > 0$, so that IP^1 corresponds to a higher level of expected profits than IP^o. Similarly, it is clear from (2) that, ceteris paribus, $\partial EU/\partial q > 0$, so that IC^1 corresponds to a higher level of expected utility than IC^o.

[25] In the analysis of reaction function equilibria, changes in exogenous variables correspond to a shift of the reaction function (e.g., Brander and Spencer, 1983, p. 709).

Figure 3

Government Intervention



that either the firm or the worker should be less cautious! Furthermore, the reader can verify that, when OSH authorities provide one party with incentives to be more careful and leave the other party's incentives unchanged (such as in the United States where safety policies are aimed mostly at firms), the economy can be pushed to a lower isoprobability locus corresponding to a higher probability of accident. This phenomenon occurs because, as mentioned by Viscusi (1979), one party can reduce its precaution level as a consequence of the increase in the other party's effort. Interestingly, such behavior could provide an intuitive reason why econometric studies have found that OSHA regulation has little or no impact on workplace safety.[26]

Altogether, it appears that safety-enhancing policies are welcome in certain circumstances. Given our assumptions, the policy prescription seems to depend on the substitutability or the complementarity of the safety inputs, something not highlighted in the previous literature.[27] Policy makers should be aware of this, and they may want to investigate the relationship between a firms' and a workers' safety inputs more closely.

The last issue to be examined is the impact of a change in legislated compensation on welfare. As argued in developing the noncooperative

[26] See Curington (1988) for a review of these studies.

[27] It should be noted that policy prescriptions may also be influenced by whether or not a Nash or a Stackelberg equilibrium is assumed.

solution, an exogenous increase in α induces the firm to increase its safety expenditure q (shifting its reaction function to the right), while it induces the worker to be less careful (shifting his or her reaction function to the left). From Figures 2 and 3, the reader can verify that, in contrast with Carmichael (1986), such a change in α does not necessarily lead, to an unambiguous improvement in welfare. An increase in α could be desirable, for instance, if unduly negligent firms induce workers to be over-prudent ($e^S > e^*$ and $q^S < q^*$), but not if $e^S < e^*$ and $q^S > q^*$. An explanation for the ambiguity of this result compared to Carmichael's is that the present model is more general than his in which the probability of an accident depends only on the firm's safety expenditure q.

Conclusion

This article has contributed to the literature by presenting a different principal-agent approach to the effect of imperfect information on the market provision of safety. Extending the analysis of Cooper and Ross (1985) to the market of workplace accidents, the problem of moral hazard was examined through the interaction of a firm and a worker in which both parties could influence the probability of an accident. Making alternative assumptions about the informational sets available to the parties, it was shown that double and single moral hazard lead to nonoptimal levels of precaution for firms and workers, and whether underprovision or overprovision of precaution was observed depended on the chosen level of insurance. Furthermore, the results of Cooper and Ross (1985) concerning the relationship between first-best and noncooperative precaution levels were shown to be sensitive to whether one of the two economic agents was risk-neutral and to the type of game equilibrium chosen.

Given the nonoptimal levels of precautions obtained in noncooperative outcomes, there appeared to be some scope for government intervention. In the context of the Stackelberg equilibrium, the welfare analysis showed that, once the compensation benefit is fixed by the state, public policy aimed at promoting safety in the workplace (by inducing economic agents to take higher levels of precaution) is not necessarily welfare-improving. In particular, it was shown that policy prescriptions were dependent on the relationship of the safety inputs: (i) when safety inputs are independent, there may be no need for imposing safety regulation on firms, but workers should be induced to be more careful, and (ii) when safety inputs are substitute, moderate safety regulation should be imposed on both parties. Furthermore, when the government influences the behavior of only one of the two parties, a policy promoting safety might not only be welfare-decreasing, but it could also fail in reaching the goal of enhancing safety.

Some extensions to the present model could be examined. First, the model could be extended to consider different accident timing or to integrate compensation benefits as a function of wages. Second, other liability rules (such as the negligence rule) in the case of an accident could be considered, and their impact as policy instruments could be analyzed. Third, it would be interesting to investigate whether or not the results presented here hold in the presence of different types of workers (adverse selection). Finally, the model could be made dynamic so that moral hazard could be mitigated with time.[28]

References

Arnott, Richard and Joseph E. Stiglitz, 1982, Equilibrium in Competitive Insurance Markets. The Welfare Economics of Moral Hazard, I: Basic Analytics, Discussion Paper No. 465 (Kingston, Ont.: Queen's University).

Boyer, Marcel and Georges Dionne, 1987, The Economics of Road Safety, *Transportation Research*, 21: 413–31.

Carmichael, Lorne, 1984, Reputations for Safety: Market Performance and Policy Remedies, Working Papers in Economics No. E-84-3 (Stanford: Hoover Institute, Stanford University).

———, 1986, Reputations for Safety: Market Performance and Policy Remedies, *Journal of Labor Economics*, 4: 458–72.

Cooper, Russell and Thomas W. Ross, 1985, Product Warranties and Double Moral Hazard, *Rand Journal of Economics*, 16: 103–13.

Curington, William P., 1988, Federal vs State Regulation: the Early Years of OSHA, *Social Science Quarterly*, 69: 341–60.

Diamond, Peter A., 1977, Insurance Theoretic Aspects of Workers' Compensation, in: Alan S. Blinder and Philip Friedman, eds., *Natural Resources, Uncertainty and General Equilibrium Systems* (New York: Academic Press).

Dionne, Georges, 1982, Moral Hazard and State-Dependent Utility Function, *Journal of Risk and Insurance*, 49: 405–22.

Emons, Winand, 1988, Warranties, Moral Hazard, and the Lemons Problem, *Journal of Economic Theory*, 46: 16–33.

Gravelle, H.S.E., 1986, Accidents, Taxes, Liability Rules and Insurance, Mimeo (London: Queen Mary College).

Holmström, Bengt, 1979, Moral Hazard and Observability, *Bell Journal of Economics*, 10: 74–91.

Kambhu, John, 1982, Optimal Product Quality under Asymmetric Information and Moral Hazard, *Bell Journal of Economics*, 13: 483–92.

Krueger, Alan B., 1988, Moral Hazard in Workers' Compensation Insurance, Mimeo (Princeton: Woodrow Wilson School, Princeton University).

Lanoie, Paul, 1989a, Occupational Safety and Health: A Problem of Double Moral Hazard, Discussion Paper no. 89–03 (Ste-Foy, Qué.: Laval University).

———, 1989b, Government Intervention in Occupational Safety and Health: a Theoretical and Empirical Analysis, Unpublished Ph.D. thesis (Kingston, Ont.: Queen's University).

[28] See, for instance, Boyer and Dionne (1987).

———, 1990, The Case of Risk Premia for Risky Jobs Revisited, *Economics Letters*, 32: 181–85.

Manga, Pran et al., 1981, *Technical Report No. 6, Occupational Health and Safety: Issues and Alternatives* (Ottawa: Economic Council of Canada, Technical Reports Series).

Mann, Duncan P. and Jennifer P. Wissink, 1988, Money-Back Contracts with Double Moral Hazard, *Rand Journal of Economics*, 19: 285–92.

Mendeloff, John, 1979, *Regulating Safety: An Economic and Political Analysis of Occupational Safety and Health Policy* (Cambridge: MIT Press).

Oi, Walter Y., 1974, On the Economics of Industrial Safety, *Law and Contemporary Problems*, 38: 669–99.

Olson, Craig A., 1981, An Analysis of Wage Differentials Received by Workers on Dangerous Jobs, *Journal of Human Resources*, 16: 165–85.

Pauly, Mark, 1974, Overprovision and Public Provision of Insurance, *Quarterly Journal of Economics*, 88: 44–62.

Rea, Samuel A. Jr., 1981, Workmen's Compensation and Occupational Safety under Imperfect Information, *American Economic Review*, 71: 80–93.

Spencer, Barbara J. and James A. Brander, 1983, International R & D Rivalry and Industrial Strategy, *Review of Economic Studies*, 50: 707–22.

Svensson, Lars E. and Assaf Razin, 1983, The Terms of Trade, Spending and the Current Account: the Harberger-Laursen-Meltzler Effect, *Journal of Political Economy*, 91: 97–125.

Thaler, Richard and Sherwin Rosen, 1976, The Value of Life Saving, in: Nestor E. Terleckyj, ed., *Household Production and Consumption* (New York: NBER Press, Columbia University).

Viscusi, W. Kip, 1978, Wealth Effects and Earning Premiums for Job Hazards, *Review of Economics and Statistics*, 60: 408–16.

———, 1979, The Impact of Occupational Safety and Health Regulation, *Bell Journal of Economics*, 10: 117–40.

———, 1980, Imperfect Job Risk Information and Optimal Workmen's Compensation Benefits, *Journal of Public Economics*, 14: 319–37.

———, 1983, *Risk per Choice: Regulating Health and Safety in the Workplace* (Cambridge: Harvard University Press).

Appendix I: Proofs

1. Proof of Proposition 1:

Proposition 1.A.: If $q^N = q^*$, then $e^N < e^*$.

1.A.1.: If $q^N = q^*$, then $e^N \nless e^*$.

Proof: Suppose $q^N = q^*$ and $e^N = e^*$. Then: **(i)** P_q^N in (10) equals P_q^* in (7), and for (10) and (7) to hold, it is necessary that $\alpha^N > \alpha^*$ since the R.H.S. (right-hand side) of (7) and (10) is identical, and since $\lambda P_q (U^a - U^n) > 0$. **(ii)** P_e^N in (9) equals P_e^* in (6), and for (9) and (6) to hold, it is necessary that $\alpha^N < \alpha^*$ since the R.H.S. of (6) and (9) is identical, and since $P_e(-\alpha)/\lambda > 0$. Recall that α is the only factor influencing the difference $U^a - U^n$. **(i)** contradicts **(ii)** proving 1.A.1.

1.A.2.: If $q^N = q^*$, then $e^N \ngtr e^*$.

Proof: Suppose $q^N = q^*$, then $e^N > e^*$. Then: **(iii)** P_q^N in (10) equals P_q^* in (7) since $P_{eq} = 0$, and for (10) and (7) to hold, it is necessary that $\alpha^N > \alpha^*$ (see (i)). **(iv)** $|P_e^N|$ in (9) is smaller than $|P_e^*|$ in (6) since $P_e < 0$ and $P_{ee} > 0$, and for (9) and (6) to hold, it is necessary that $\alpha^N > \alpha^*$ (see (ii)). **(iii)** contradicts **(iv)** proving 1.A.2.

Q.E.D.

Propositions 1.B. and 1.C.: The proofs are similar to the preceding ones and are omitted.

2. Proof of Proposition 2:

Proposition 2.C.: $q^N < q^*$ and $e^N < e^*$ when $\alpha^N = \alpha^*$.

2.C.1.: When $\alpha^N = \alpha^*$, $e^N < e^*$.

Proof: When $\alpha^N = \alpha^*$, for (6) and (9) to hold, $|P_e^N|$ must be larger than $|P_e^*|$ since $P_e < 0$ and $P_{ee} > 0$, and this implies that $e^N < e^*$.

2.C.2.: When $\alpha^N = \alpha^*$, $q^N < q^*$.

Proof: When $\alpha^N = \alpha^*$, for (7) and (10) to hold, $|P_q^N|$ must be larger than $|P_q^*|$ since $P_q < 0$ and $P_{qq} > 0$, and this implies that $q^N < q^*$.

Q.E.D.

Propositions 2.A. and 2.B.: They follow directly from Proposition 2.C. and from the fact that $dq/d\alpha > 0$ (see (10)) and that $de/d\alpha < 0$ (see (9)).

3. Proof of Proposition 3: When $\alpha^s = \alpha^*$, then $q^s > q^*$ and $e^s < e^*$.

3.A.: When $\alpha^S = \alpha^*$, then $q^S > q^*$.

- From (5) and (II.6), it follows that $\lambda^s = \lambda + \Psi$ where Ψ has a positive value.
- Therefore, for (7) and (II.7) to hold, it is necessary that $q^S > q^*$ since $P_q < 0$, $P_{qq} > 0$, $\lambda^S > \lambda$ and $\alpha^S = \alpha^*$.

3.B.: When $\alpha^S = \alpha^*$, then $e^s < e^*$.

- Same proof as the preceding one.

Q.E.D.

4. Proof of Proposition 4:

4.A.: The locus depicting the first-best equation (6) is everywhere on the right of the locus depicting (9) in e,q space.

Proof:

- Both equations (9) and (6) describe a locus in e,q space.
- Equation (6) can be rewritten as: $P_e(U^a - U^n) - 1 = P_e(\alpha)/\lambda$ where the R.H.S. is negative.
- Equation (9) implies that, at given q, when one moves away from the Stackelberg equilibrium by increasing e, $P_e(U^a - U^n)$ becomes smaller since $P_{ee} > 0$ and since $\alpha^S = \alpha^*$, so that $P_e(U^a - U^n) - 1$ becomes negative.
- Therefore, since the R.H.S. of (6) is negative, it implies that the locus depicting (6) is everywhere on the right of the locus depicting (9) in e,q space.

4.B.: The locus depicting (7) is everywhere above the locus depicting equation (II.7) in e,q space.

Proof: When $P_{eq} > 0$, there is term $P_e(-\alpha)de/dq$ added on the L.H.S. of equation (II.7) in Appendix II.

- Both equations (II.7) and (7) describe a locus in e,q space.
- Equation (II.7) can be rewritten as $P_q(U^a - U^n)\lambda^S + P_q(-\alpha) - 1 = P_e(\alpha)de/dq$ with R.H.S. positive when $P_{eq} > 0$.
- From (5) and (II.6), it follows that $\lambda^s = \lambda + \psi$ where ψ has a positive value. Therefore, (7) can be rewritten as $P_q(U^a - U^n)\lambda^s + P_q(-\alpha) - 1 = -\psi P_q(U^a - U^n)$ where the R.H.S. is negative.
- Equation (II.7) implies that, at given e, when one moves away from the Stackelberg equilibrium by

increasing q, P_q $(U^a - U^n)\lambda + P_q(-\alpha) - 1$ becomes lower since $P_{qq} > 0$ and since $\alpha^S = \alpha^*$.

- Therefore, since the R.H.S. of (7) is negative, it implies that the locus depicting (7) is everywhere above (II.7) in e, q space.

Q.E.D.

Appendix II

II.A. The First Stage of the Nash Game:

The Lagrangian is:

(II.1) $$\underset{w,\alpha,\lambda^N}{L} = f - w - q(w,\alpha) - P(e(w,\alpha),q(w,\alpha))\alpha$$
$$+ \lambda^N\{P(e(w,\alpha),q(w,\alpha))U^a + [1 - P(e(w,\alpha),q(w,\alpha))]U^n - e(w,\alpha) - \overline{U}\}$$

The F.O.C. with respect to w and α after simplification (using (9) and (10)) are:

(II.2) $$w: -1 + \lambda^N U_w = 0$$

(II.3) $$\alpha: -P + \lambda^N P U^a_\alpha + P_e(-\alpha)\frac{de}{d\alpha} + \lambda^N P_q(U^a - U^n)\frac{dq}{d\alpha} = 0;\ \lambda^N > 0$$

II.B. The First Stage of the Stackelberg Game:

(II.4) $$\underset{w,\alpha,q,\lambda^S}{L} = f - w - q - P(e(q,w,\alpha),q)\alpha)$$
$$+ \lambda^S\{P(e(q,w,\alpha),q)U^a + [1 - P(e(q,w,\alpha),q)]U^n - e(q,w,\alpha) - \overline{U}\}$$

The F.O.C. with respect to w, α and q are:

(II.5) $$w: -1 + \lambda^S U_w = 0$$

(II.6) $$\alpha: -P + \lambda^S P U^a_\alpha + P_e(-\alpha)\frac{de}{d\alpha} = 0$$

(II.7) $$q: P_q(U^a - U^n)\lambda^S + P_q(-\alpha) - 1 = 0$$