



# Working Papers

[www.cesifo.org/wp](http://www.cesifo.org/wp)

## The Value of Technology Improvements in Games with Externalities: A Fresh Look at Offsetting Behavior

Michael Hoy  
Mattias K. Polborn

CESIFO WORKING PAPER NO. 4798

CATEGORY 1: PUBLIC FINANCE

MAY 2014

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

CESifo

Center for Economic Studies & Ifo Institute

# The Value of Technology Improvements in Games with Externalities: A Fresh Look at Offsetting Behavior

## Abstract

How should we evaluate the welfare implications of improvements to safety technologies in the presence of offsetting behavior? We model this problem as a symmetric game in which each player's payoff depends on his own action and the average action of the other players, and analyze under which conditions an improved technology increases or decreases both the level of precautionary activity and equilibrium utility of players. For mandatory safety technologies, the direction of the welfare effect depends on whether the externality between players is positive or negative, and on whether the improved technology increases the individually optimal activity level, taking the activity of other players as given. For safety technologies that individuals can choose whether to employ, we show that an individual will generally expend too much on reducing the size of loss but, depending on conditions that we specify, either too much or too little on features that reduce the individual's probability of loss.

JEL-Code: C700, D600, D700, D800.

*Michael Hoy*  
*Economics Department*  
*University of Guelph*  
*Canada - N1G 2W1 Guelph, Ontario*  
*mhoy@uoguelph.ca*

*Mattias K. Polborn*  
*Department of Economics*  
*University of Illinois*  
*410 David Kinley Hall, 1407 W. Gregory*  
*USA - 61801-3606 Champaign, IL*  
*polborn@uiuc.edu*

April 24, 2014

# 1 Introduction

The seminal contribution of Peltzman (1975) addressed the possibility that legally mandated improvements to safety devices on automobiles may not lead to anticipated reductions in the overall accident rate. He proposed, and tested for, the possibility that driver responses to the mandated ‘improved safety environment’ (seat belt use) could lead to offsetting effects in the form of reduced care in driving habits. His important insight is that such behavioral change, which reflects an individually rational response to the lowered cost or severity of accidents due to the mandated use of safety belts, could, from a societal perspective, be entirely offset by the induced behavioral change. The paper spearheaded a wide range of applications of the offsetting hypothesis for numerous phenomena including traffic accidents, workplace accidents, food safety, rates of HIV (e.g., Geoffard and Philipson (1996)) and other infectious diseases, and many more.

Peltzman relied on a reduced form model to develop his empirical analysis. Others have subsequently developed explicit behavioral models to investigate the implications of offsetting behavior for a wide range of applications. A vast number of papers have also empirically investigated the prevalence and magnitude of offsetting behavioral effects (see Harless and Hoffer (2003) and references therein). Our goal is to develop a theoretical approach that will provide insights into how the broad range of phenomena with offsetting behavior should be modeled. Although addressing a rather different phenomenon, the paper by Ehrlich and Becker (1972) informs our modeling approach. They consider the choice of both self protection, self-insurance and market insurance to develop insights into the substitutability and complementarity of choices that people can take to reduce the consequences of accidents or other bad outcomes. Although we focus on nonfinancial losses in our model, we consider both changes in the magnitude or severity of loss as well as the choice of level of precautionary activity that influences the probability of a loss and how these relate. Besides the introduction of safety technologies, such as seat belts or airbags, which reduce the magnitude of a loss, we also consider safety technologies which, conditional on a given level of precautionary activity, affect the probability of a loss occurring, such as improved brake systems (e.g., ABS) on automobiles. The former we call an improved loss mitigation technology (LMT) while the latter we call an improved probability reduction technology (PRT). We show that an improved LMT will always lead to a reduction in an individual’s optimal level of precautionary activity, consistent with Peltzman’s hypothesis. We also develop conditions under which this effect is sufficiently strong that introduction of an improved LMT leads to a reduction in utility (welfare) even when there is no direct cost to the improved technology.

Whether an improved PRT, such as more effective brakes for an automobile or rumble strips on the highway, leads to a lower or higher accident rate depends not only on the di-

rect effect of the technology, but also on whether the new technology increases or decreases the equilibrium level of precaution. That effect depends on how the technology affects the marginal effectiveness of precaution, and can go in either direction. An improved braking system may increase the marginal benefit of attentiveness for avoiding accidents, while rumble strips can be expected to reduce the individual's perceived value of more frequent rests while driving.

Again, we also determine conditions under which PRT improvements lead to a worsening or improvement in welfare. If, for example the adoption of ABS leads to increased marginal effectiveness of personal precautionary efforts, then this would not only lead to a decrease in the equilibrium rate of accidents. The equilibrium value of mandatory use of this intervention would exceed the computation based on no behavioral effects. This result is opposite to that implied by improvements to LMT, such as analyzed by Peltzman.

We model explicitly how the probability that an individual experiences a loss depends both on the individual's own level of care and on the average level of care of other agents. This externality interacts with the effect of the improved safety technology on individuals' choice of precaution to determine the equilibrium effect on both the accident rate and the impact on welfare. We show how the assessment of the social value of improved PRTs and LRTs are determined by the interaction of these effects.<sup>1</sup> First, we provide a general model that captures or subsumes many of the relevant applications of offsetting behavior. The usual application of the offsetting hypothesis implies that the value of an 'improvement' to safety is lowered by individually rational reactions to such changes. We characterize conditions under which the possibility that behavior changes as a result of the new technology leads to an increased or decreased level of welfare relative to the case without behavioral effects having been taken into account. The decisive condition turns out to be whether the level of care that individuals exert, measured by how costly it is for individuals to be careful, increases or decreases. We then develop more specific models in order to offer more detailed insights into specific phenomena in which offsetting behavior is an important consideration.

A few remarks here will help to explain how our analysis fits into the existing literature, most specifically that on road safety and control of infectious diseases. When we refer to the individual's level of precaution we mean things such as attentiveness to road hazards while driving or use of safer sex practices. These are assumed unobservable to the social planner (or government). Our analysis is designed to consider how such choices create

---

<sup>1</sup>Of course, some innovations can affect both the probability and magnitude of loss. A good example is the use of ABS which can lower a person's likelihood of impact with another car and also the effect of such impacts that aren't avoided. In a somewhat more complicated manner, improved antiviral drugs for HIV decrease the communicability of disease from others while also reducing severity of loss for those who do become infected.

externalities for others<sup>2</sup> and how these choices are affected by other safety measures, whether these are observable features purchased by individuals themselves or are publicly provided (or mandated).

We do not consider many of the indirect measures used for imperfectly observing (and controlling) such individual choices of precaution. These include experience rating by insurers, liability through negligence rules enforced through the legal system, or imperfect monitoring such as police enforcement of traffic regulations that others have studied (e.g., Boyer and Dionne, 1987). Our reason is twofold: (1) in some instances such instruments are very costly or impossible to effect and (2) we wish to emphasize in isolation of other issues how the externality created by moral hazard that affects other agents impacts on social welfare. We recognize that, even without this direct type of externality, individual moral hazard can create a negative externality through an insurance pool (e.g., see Gossner and Picard, 2005). We leave aside these sorts of issues in this paper, although they are all well worth exploring in future work. Our main contribution here is to demonstrate how modeling offsetting behavior as a problem involving moral hazard in teams offers new insights into the phenomenon and points to directions for future research.

In the next section we develop a unified model for analyzing both the positive and normative implications of offsetting behaviour. In the context of this model, we review the related literature in the following section. In section 4 we develop applications of the unified model with separate functions to describe the probability of loss and the size of loss. Both cases in which the level of safety technology is exogenously imposed or endogenously chosen are considered. This allows us to further develop the intuition of the simple unified model and provide useful guidance about modeling other specific economic settings in which offsetting behaviour is an important element. The final section concludes and offers suggestions for further research on this topic.

## 2 The unified model

We begin with a model that unifies many existing models that highlight offsetting behavior. Consider a game between a continuum of players. All players are symmetric and each player's payoff depends on his own activity level  $x$ , the average activity level  $\bar{x}$  of the other players, and  $\theta$  which parameterizes the technology. (Note: We will often refer to  $x$  as a level of precaution.) Each player's problem therefore is

$$\max_x f(x, \bar{x}, \theta). \tag{1}$$

---

<sup>2</sup>Our paper is essentially an application of the phenomenon of *moral hazard in teams*. See Holmstrom, 1982) for a general characterization of this problem and Cooper and Ross (1985), Lanoie (1991), Pedersen (2003), and Risa (1992, 1995) for specific applications.

We assume that  $f$  is increasing in its third argument: Holding all players' actions fixed, a player's payoff increases as technology improves. As to the second argument, we allow for both positive ( $f_2 > 0$ ) and negative ( $f_2 < 0$ ) externalities from other players' actions. Finally, we assume that  $f$  is strictly concave in its first argument, so that the first order condition

$$\frac{\partial f}{\partial x} \equiv f_1 = 0 \quad (2)$$

is necessary and sufficient for a global optimum. In general, we denote partial derivatives with respect to the  $i$ th argument as  $f_i$ .

Totally differentiating (2), we find that the change of the equilibrium level of  $x$  is given by

$$\frac{dx}{d\theta} = -\frac{f_{13}}{f_{11} + f_{12}}, \quad (3)$$

For the game to have a stable equilibrium, we assume that, when the average action of the other players increases by 1, the individually optimal action increases by less than 1. Formally, we have the following:

**Assumption 1.** *Let  $\bar{x}(\theta)$  be the equilibrium level of  $x$ , given technology  $\theta$ . Then we have  $f_{12}(\bar{x}(\theta), \bar{x}(\theta), \theta) < -f_{11}(\bar{x}(\theta), \bar{x}(\theta), \theta)$  for all  $\theta$ .*

Given Assumption 1, the sign of  $dx/d\theta$  in (3) is the same as the sign of  $f_{13}$ . Intuitively, if an increase in  $\theta$  increases the marginal payoff to  $x$  (i.e., if  $f_{13} > 0$ , then the equilibrium level of  $x$  increases and vice versa.

Let us now turn to the effect of a technological improvement on the equilibrium utility of players,

$$V(\theta) = f(\bar{x}(\theta), \bar{x}(\theta), \theta). \quad (4)$$

Using the envelope theorem, a change of  $\theta$  has the following effect on equilibrium utility:

$$V'(\theta) = \frac{\partial f}{\partial \theta} + \frac{\partial f}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial \theta} = f_3 - f_2 \frac{f_{13}}{f_{11} + f_{12}}. \quad (5)$$

In many applications, we can normalize  $\theta$  such that  $f_3 > 0$ . That is, increasing  $\theta$  means progress: Holding constant the actions of all players, each player's payoff increases. The second term measures the effect of strategic interaction between players. If the externality between players is positive ( $f_2 > 0$ ), then the sign of the interaction effect is the same as the sign of  $f_{13}$ : If  $f_{13} > 0$ , then the equilibrium level of  $x$  increases, which has additional beneficial effects on the equilibrium utility of all players. If, instead,  $f_{13} < 0$ , then the equilibrium level of  $x$  decreases, thereby counteracting the positive direct technology effect. In the examples in the next section, we show that it is even possible that the overall effect of technical progress on players' equilibrium utility is negative.

In case of a negative externality ( $f_2 < 0$ ), these effects go in the other direction. If  $f_{13} > 0$ , then the equilibrium level of  $x$  goes up which creates more negative externalities

and hence counteracts the direct effect of technical progress. If  $f_{13} < 0$ , then the equilibrium level of  $x$  goes down, reducing the negative externality between players and further increasing their equilibrium utility level. Hence we have

**Proposition 1.** *The following are sufficient conditions for technical progress to increase the equilibrium utility level of players:*

1. *The externality between players is positive ( $f_2 > 0$ ) and technical progress increases the equilibrium activity level ( $f_{13} > 0$ ).*
2. *The externality between players is negative ( $f_2 < 0$ ) and technical progress decreases the equilibrium activity level ( $f_{13} < 0$ ).*
3. *The externality between players is zero.*

*If neither condition holds, then the equilibrium utility change is smaller than the direct effect of technical progress (holding the equilibrium level of  $x$  fixed). In particular, the overall effect of technical progress on players' equilibrium utility level is negative if*

$$f_3(f_{11} + f_{12}) - f_2f_{13} > 0. \quad (6)$$

Consider a model of traffic safety with the probability of an individual experiencing an accident decreasing in both her own and in others' level of precaution. This means that the externality is positive ( $f_2 > 0$ ); i.e., an individual benefits from others increasing their level of precaution. Since the magnitude of the loss associated with an accident is reduced with an improvement in the LMT ( $\theta$ ), this decreases the marginal benefit of precaution for all individuals ( $f_{13} < 0$ ). The introduction of mandatory seat belts or airbags represents such a case. Therefore, no set of sufficient conditions in Proposition 1 holds and so equation (6) applies (recall that  $f_{11} + f_{12} < 0$  due to Assumption 1). The larger is the term  $f_2f_{13}$ , the more likely will condition (6) hold and so the improvement in the LMT leads to a reduction in welfare. We explore this possibility in the applications section below.

Conditional on all individuals maintaining their level of precaution, many technological advancements such as antilock brakes, rumble strips on the edge of roadways or drowsiness alert systems reduce the probability of an accident. We refer to these as probability reduction technologies (PRT).

PRTs may increase or decrease the marginal benefit of an individual's own level of precaution. It seems plausible that better braking systems enhance an individual's own efforts to avoid an accident, implying  $f_{13} > 0$ , while rumble strips and drowsiness alert systems represent substitutes for taking care to be well rested and alert while driving (i.e., taking appropriate rest breaks on a long trip) and so imply  $f_{13} < 0$ . Thus, some

improvements to PRTs such as the first example above fall into the category of case 1 in Proposition 1, while others, such as the second and third examples above, require evaluation of condition (6) in order to determine whether an improvement will lead to an increase or decrease in welfare.

While payoffs in our model are assumed to depend on the average action level of the other players, one can also interpret the simple model in this section as a two person game in which the “actual opponent” is randomly drawn from the large set of other players. For example, modeling the accident probability for individuals traveling on a two lane roadway may also be well suited to a series of pairwise interactions. To account for such scenarios, one may simply replace the pair  $\{x, \bar{x}\}$  in equation (1) with  $\{x_1, x_2\}$ ; where  $x_i$  is the level of activity (precaution) of individual  $i = 1, 2$ . The model now becomes a standard Cournot-Nash equilibrium model.<sup>3</sup>

The modification of our simple model based on the average level of precaution to a two person game highlights an important assumption that is implicit in our original model. Suppose we consider a model with a finite (but possibly large) number of individuals indexed  $i = 1, 2, \dots, n$ . Then if we model player  $j$ 's payoff or utility as depending on his own level of precaution,  $x_j$ , the vector of all others level of precaution,  $\mathbf{x}_{-j} = (x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ , and the technology parameter,  $\theta$ , then player  $j$ 's utility function becomes:

$$f^j(x_j, \mathbf{x}_{-j}, \theta) \tag{7}$$

and there are  $n-1$  cross partials  $f_{ji}^j, i \in \{x_1, x_2, \dots, x_{j-1}, x_{j+1}, \dots, x_n\}$  to represent possible external effects from others level of precaution. In most of our applications, all of these partial derivatives are positive and the sign of  $f_{j\theta}^j$  depends on the context of the application (as above). With homogeneous individuals and a focus on symmetric equilibrium, the simplification to our model, equation (1), is reasonable.

### 3 Related Literature

Much of the literature about the phenomenon of offsetting behaviour has been directed at determining empirically its size in a wide variety of economic settings. Following Peltzman (1975), the most commonly investigated phenomenon is road safety. Various studies have, for example, attempted to determine if adoption of new technologies such as airbags or antilock braking systems lead to a reduction in driver safety through increased speeds (see, e.g., Winston, et al. (2006) and Harless and Hoffer (2003)). Other papers empiri-

---

<sup>3</sup>Alternatively, one can think of sexual relationships that may result in a sexually transmitted disease (STD). In such an application,  $x_i$  could represent condom use in a current or previous sexual encounter, frequency of past unprotected and/or protected sexual encounters, testing for STDs, etc.



cally investigate phenomena related to workplace safety (e.g., Lanoie (1992)), sports (e.g., Potter (2011) on formula 1 racing and McCannon (2011) on basketball), food safety (e.g., Miljkovic (2011), et al.), and even policies aimed to foster healthy living choices (e.g., substitution between taxed soft drinks such as sodas, and other high calorie drinks (Fletcher, et al. (2010))).

There is also a broad range of theoretical models of offsetting behaviour in a wide variety of contexts. In fact, one of the challenges in developing a unified model is due to the wide range of applications. This is also a reason why such an effort is valuable as our model can be employed as a starting point for any such study. Our unified model stresses the strategic aspect that a model should possess as well as a formal externality effect. Two important elements of the strategic aspect are the manner in which the technology affects individual incentives (through the term  $f_{13}$ ), as well as the impact that a change in others' behaviour has on any individual's own behaviour (through the term  $f_{12}$ ). One must also understand the strength of the externality effect of others' behaviour on one's own well-being or utility (through the term  $f_2$ ).

Some of these elements are present in existing theoretical models of offsetting behaviour, but usually one or more of them is missing, or the models that capture all these aspects typically relate to a very specific environment and so are not easily transferable to other applications. We review here some such theoretical models.

Kunreuther and Heal (2003) provide a very interesting application of interactions and how these influence collective safety outcomes. They investigate, as an example, how an airline's decision to adopt a security system for baggage checking is affected by the security decision of other airlines. Any airline may choose a baggage checking method that influences the security of its flights arising from passengers with trips originating on its own flights. The externality arises due to the acceptance by each airline of the baggage of travelers who are transferring from other airlines making connecting flights when those passengers' bags are not rechecked, as is typical. If the other airlines do not adopt a secure baggage checking system, then this compromises the ability of any given airline to avoid bad outcomes. This creates a complementarity between airlines' decisions such as the interaction effect that is embodied by the term  $f_{12}$  in our model (with  $f_{12} > 0$  in their case). The goal of their paper, however, differs from ours as their model is designed to understand how one might improve overall welfare through influencing or managing individual decisions in the context of the interaction effects. They do not consider the relationship between improved technologies as a second dimension of effort which is at the heart of our model (i.e., the interaction of the variables  $\theta$ , which represents an improvement in a safety technology of either an LMT or PRT type, with individuals' decisions on effort, characterized by the variable or activity level  $x$ ).

In a paper with some of the same elements as Kunreuther and Heal (2003), Muermann and Kunreuther (2008) model a problem of optimal investment in self-protection of insured individuals in the presence of interdependencies with others' decisions (i.e.; "contamination"). In their model, an individual's own effort affects his probability of an accident, which we will refer to as due to his own fault. If an agent avoids an "own fault" accident, there is a possibility that he will be involved in an accident that occurs due to the fault of some other person, with a probability related to the other person's effort choice. This is the source of contamination or externality. It turns out that each agent's level of effort affects the marginal benefit of the other agent's effort choice, but in a very specific manner. It turns out that both partial insurance and use of an "at-fault", as opposed to a "no-fault", insurance system can lead to a better set of decisions (i.e., which improve everyone's utility).

Another paper that has some of the same properties and objectives as our paper is that of Gossner and Picard (2005). Their goal is to investigate how to value the benefit of an improvement in road safety in the presence of an offsetting effect. The loss in their model is financial and the source of externalities is through the insurance market rather than through some explicit interaction effects as reflected by the contamination effects of the papers discussed above or through the term  $f_2$  in our model. They do, however, consider a similar problem as in our paper by taking into account how changes in road safety affect precautionary effort levels of individuals. Due to the fact that losses are financial, they also investigate the implications of drivers' risk aversion on the value of improvements to road safety.

Besides the sort of models developed for specific contexts that are discussed above, there are some general models that have been proposed to study the phenomenon of offsetting behaviour. Our model should be thought of as further developing those models. Of particular relevance is the paper by Blomquist (1986). He develops a general model of driver safety behaviour and demonstrates the positive result that "under plausible conditions a change in exogenous safety, which is beyond driver control, causes a compensatory change in driver effort in the opposite direction" ... (Blomquist, 1986, p. 371). His model has both dimensions of safety as does ours (i.e., exogenous safety factors -  $\theta$  in our model - and endogenous driver safety choice -  $x$  in our model) and provides a useful comparative static result describing conditions under which the choice of exogenous safety may reduce the driver's own effort to avoid bad outcomes. However, he does not explicitly model an externality factor or strategic behaviour. Nor does he address welfare implications.

Neill (1993) also develops a model to determine conditions under which the probability of an accident increases or decreases as a result of an increase in the level of an imposed safety technology or regulation. As in our model, this depends on how the increase in

the imposed safety technology affects the marginal benefit of individuals' own levels of precaution. His paper investigates how this relationship between the safety technology and the individual's effort to avoid accidents impact on the choice of self-insurance (LMT in our terminology and safety devices in his). However, he does not address the normative implications of imposed safety technologies.

Hause (2006) also develops a general model of the offsetting phenomenon. He points out (pp. 689-690) that "Despite accumulating evidence on the empirical relevance of OB (offsetting behaviour), none of the theoretical literature has provided a model determining formal conditions under which dominant or partial OB occurs, much less the magnitude of the OB effect on expected accident loss". By a dominant effect Hause means that the OB effect (change in own effort of accident avoidance) results in no net change in the expected accident loss. By a partial OB effect is meant that the net effect of the safety regulation or technology is a reduction in the net expected accident loss, but less than the direct effect. We show, however, that knowing the extent or size of the offsetting effect does not allow one to conclude whether the imposed safety technology improves or worsens welfare, as is clear from our Proposition 1 and examples in section 4.1. His model includes a level of safety regulation (imposed technology) as well as an individually chosen level of accident avoidance, but does not include an explicit externality effect or interaction effect of others' safety choice with an individual's own safety choice.

The most important advantage of our model is that we combine the elements of an explicit treatment of (1) strategic behaviour, including consideration of whether activity levels are strategic complements or substitutes, (2) an externality effect between people's activity levels, (3) how the safety technology affects the marginal value of precaution, and (4) whether the adopted safety technology is an LMT (mitigates loss) or a PRT (reduces probability of loss). Unlike most earlier models, we not only consider the positive implications of offsetting behaviour (i.e., the extent to which OB offsets the direct reduction in exposure to the loss due to the imposed safety technology) but also develop conditions for whether the net effect on welfare is positive or negative. We develop these insights further in the next section. We then consider scenarios in which individuals purchase a given level of safety technology rather than have one imposed on them.

## 4 Applications

### 4.1 Exogenous Safety Improvements

Consider a society with many individuals who are all subject to the possibility of suffering a loss  $L(\lambda)$ . As in the motivating examples described in the introduction,  $L$  can be thought

of as a result of catching a disease or having a car accident. The parameter  $\lambda$  describes the state of the loss mitigation technology that influences how severe a loss is if it occurs. One can think of a treatment for a disease that improves the quality of life of affected patients, or a car safety device like an airbag that mitigates the health consequences of an accident. We assume that  $L' < 0$ .

In the beginning of the game, players all simultaneously choose their level of precaution,  $p$ . The probability that individual  $i$  suffers a loss is given by the function  $D(p_i, \bar{p}, \theta)$  that depends on the level of care of individual  $i$ ,  $p_i$ , the average level of care that everybody else in the society exerts,  $\bar{p}$ , and the PRT technology  $\theta$  that parameterizes the function  $D$ .

We assume that all first derivatives of the function  $D$  are continuous, and have the following signs: First,  $D_1 \equiv \frac{\partial D}{\partial p_i} < 0$ , so that a higher level of care of individual  $i$  strictly decreases his probability of suffering a loss. Furthermore,  $D_{11} > 0$ , so that there are decreasing returns to scale in the level of care. Second,  $D_2 \equiv \frac{\partial D}{\partial \bar{p}} \leq 0$ , so that a higher level of care by everybody else also (at least weakly) diminishes an individual's loss probability. A strictly negative derivative is intuitively plausible for the traffic application as well as for communicable diseases, while other diseases like cancer or heart attacks would have a zero derivative in the second argument. Third, we parameterize the state of knowledge  $\theta$  such that a higher level of  $\theta$  corresponds to a smaller loss probability for all levels of precaution:  $D_3 \equiv \frac{\partial D}{\partial \theta} \leq 0$ . Examples of  $\theta$  would be an anti-lock brake system (ABS) or new information about how diet influences the probability of getting some type of cancer.<sup>4</sup>

Being more careful involves a direct utility cost for the individual, given by the function  $c(p_i)$ . We assume the usual conditions that guarantee an interior optimum for  $p_i$ :  $c'(p_i) > 0$  and  $c''(p_i) > 0$  for all  $p_i$  and  $c'(0) = 0$ .

Individual  $i$ 's objective is to minimize the sum of the expected loss and the cost of care,  $D(p_i, \bar{p}, \theta)L(\lambda) + c(p_i)$ , so we can recast the problem to be consistent with equation (1) as one of choosing  $p_i$  to maximize:

$$V(\lambda, \theta) = -D(p_i, \bar{p}, \theta)L(\lambda) - c(p_i). \quad (8)$$

The first order condition corresponding to this problem is

$$-D_1(p_i, \bar{p}, \theta)L(\lambda) - c'(p_i) = 0. \quad (9)$$

---

<sup>4</sup>Whenever new information shows that previously available actions have precautionary benefits (such as that taking a low dose of aspirin each day reduces the likelihood of heart attacks), revealed preference arguments suggest that implementing the precautionary action likely has some positive cost for individuals, because otherwise, they would have taken the action already before knowing about its precautionary value. However, if an individual chooses to implement the precautionary action, then revealed preference arguments imply that the net benefit is positive. Our functions should be understood as measuring this net benefit.

Since  $D_{11}$  and  $c''$  are assumed to be positive, this first order condition is also a sufficient condition for maximizing the individual's utility.

Applying Proposition 1, we have that

$$\frac{dV}{d\theta_1} = -DL' - D_2L \frac{-DL'}{(D_{11} + D_{12})L} = -DL' + D_2 \frac{DL'}{(D_{11} + D_{12})}. \quad (10)$$

The first term is the direct effect of reduced loss severity on expected utility and so is positive ( $L' < 0$  by definition of a LMT improvement). The second term is the indirect effect that takes into account the effect of the equilibrium response: since an accident is less harmful, players reduce their level of care, which leads to an increase in the negative externality from this reduced level of care. The total effect of a reduced loss size on expected utility is therefore uncertain and depends on which of the two countervailing effects dominates.

As for changes in  $\theta$ , we have

$$\frac{dV}{d\theta} = -\frac{\partial D}{\partial \theta}L - D_2L \frac{D_{13}L}{(D_{11} + D_{12})L}. \quad (11)$$

Again, the direct effect in the first term is clearly positive ( $\frac{\partial D}{\partial \theta} < 0$  by definition of a PRT improvement). The indirect effect is positive if  $D_{13} > 0$ , because then an increase in  $\theta$  increases the equilibrium level of care, and consequently the negative externality decreases. If  $D_{13} < 0$ , then an increase in  $\theta$  decreases the equilibrium level of precaution, which increases the negative externality. As a consequence, the total effect of an increase of  $\theta$  is ambiguous if  $D_{13} < 0$ , while it is definitely positive if  $D_{13} > 0$ .

We now show by example that an increase in  $\lambda$ , resulting in a decreased potential loss  $L$ , can be counter-productive from a welfare point of view; that is, a worsening of the negative externality may overwhelm the positive direct effect of the improved loss mitigation technology.

Suppose that the probability of a loss is given by

$$D(p, \bar{p}) = \kappa - bp - (1 - b)\bar{p}, \quad (12)$$

so that  $b$  (with  $0 < b < 1$ ) measures the extent to which the loss probability is influenced by the individual's own level of care, while  $(1 - b)$  measures the extent that this probability is influenced by the average level of care of other players.<sup>5</sup> Smaller  $b$  implies own effort is less important relative to others effort.

---

<sup>5</sup>One could use unrelated parameters, say  $\beta$  (with  $0 < \beta < 1$ ), rather than  $(1 - b)$ , to measure the effect of the average level of care of other players. However, our single parameter approach is mathematically convenient and  $b$  measures the relative importance of own effort versus others effort (the externality effect) in determining the probability of loss.

Furthermore, assume that the cost of care is given by  $c(p) = \alpha p^2$ . The individual then chooses  $p$  to minimize

$$[\kappa - bp - (1 - b)\bar{p}]L + \alpha p^2,$$

which yields  $p = \frac{bL}{2\alpha}$  for an interior solution, and, of course, this is also the average level of care of everybody else in equilibrium. Note that, in order for an interior solution to apply, we have to restrict parameters such that the probability of a loss is non-negative and less than one in equilibrium:  $\frac{bL}{2\alpha} < \kappa < 1 + \frac{bL}{2\alpha}$ .

Substituting the optimal value for both  $p$  and  $\bar{p}$  in (12) above yields that the equilibrium total cost  $T$  (i.e., expected loss plus cost of care) is

$$T = \left[ \kappa - \frac{bL}{2\alpha} \right] L + \alpha \left( \frac{bL}{2\alpha} \right)^2. \quad (13)$$

Differentiating with respect to  $L$  gives

$$\frac{dT}{dL} = \kappa - \frac{bL}{2\alpha}(2 - b). \quad (14)$$

If  $b = 1$ , i.e. the probability of a loss depends only on the individual's own behavior, then a decrease in  $L$  always helps the individual, as the expression in (14) becomes  $\kappa - \frac{L}{2\alpha} > 0$ . (The sign follows from the assumption above that  $\kappa > \frac{bL}{2\alpha}$ ). This result is intuitive, since  $b = 1$  implies that there are no externalities and thus each individual is not negatively affected by the fact that all other individuals behave less carefully when  $L$  decreases.

However, this result changes when  $b$  is small and  $\kappa$  is not much larger than  $\frac{bL}{2\alpha}$ ; that is, own effort is unimportant relative to the effort of others in determining the loss probability. Then, the derivative in (14) is negative, which means that a decrease in  $L$  actually *increases* the total cost  $T$  to individuals.

Intuitively, a decrease in  $L$  induces all individuals to reduce their level of care. Just by itself (keeping other people's level of care constant), this is a beneficial effect for the individual, as his cost of care decreases by more than the expected cost of a loss increase. However, since the level of care of the other individuals is not fixed but also decreases in equilibrium, an additional negative externality is generated by the decrease in  $L$ , which outweighs the positive direct effect.

Some important policy considerations can be brought out by applying a few specific sets of parameters to this very simple model. Suppose  $L$  represents quality adjusted life years for a person who incurs a disease and  $c(p) = \alpha p^2$  is the individual's subjective cost of precaution level  $p$ . Consistency requires that  $c(p)$  be measured in units equivalent to quality adjusted life years (e.g., an increase in  $c(p)$  of one unit is equivalent to giving up one quality adjusted life year).

The individual's objective is then to choose  $p$  to maximize the expected value of quality adjusted life years (inclusive of cost of precaution) or, equivalently, to minimize  $T$ . An

interesting question is how one should value the effect of improving treatment should disease occur (i.e., a reduction in  $L$ ) for scenarios which vary according to the extent of offsetting behavior.

We consider a set of parameterized examples involving the reduction in  $L$  from a level of  $L = 10$  to  $L = 9$ . A naive perspective might be that, given the existing probability of disease associated with the level  $L = 10$ , the value of an improvement in treatment of one unit is simply  $D^0 \cdot 10 - D^0 \cdot 9 = D^0$ , where  $D^0$  is the equilibrium value of the (initial) probability of disease when  $L = 10$ . One might argue that any offsetting behavior should reduce this value since the probability of disease increases as a result of the improvement in quality adjusted life years. In fact, the opposite is true when there is no externality effect since the reduction in cost of prevention due to the reduction in  $p$  that is induced by improved treatment is the result of the individual's optimization behavior and this increase in expected utility is higher than if he were restricted to maintain  $p$  (and hence  $D^0$ ) at its previous level.

If there is a negative externality associated with this reduced prevention effort, then it is certainly possible that the value of the improvement to quality of treatment is less than  $D^0$  and even negative (as discussed above). However, it is still possible that the value of the improvement in treatment can exceed  $D^0$ , even in the presence of a negative externality effect. This can happen if the private gain to everyone from reducing  $p$  exceeds the loss to everyone due to the negative externality associated with this offsetting behavior. Moreover, the cases described below for this example demonstrate that even for 'quite large' offsetting effects (i.e., large increases in  $D$  induced by the reduction in  $L$ ) it is possible that the naive approach undervalues, rather than overvalues, the improved treatment.

All cases in Table 1 involve a starting value of  $L = 10$ , and a reduction of 10% to  $L = 9$ . Our first case involves no externality (i.e.,  $b = 1$ ). For  $\kappa = 1.2$ , we find that the optimal level of precaution is  $p = 1$  which implies (initial) loss probability  $D^0 = 0.2$ , expected loss  $D^0 \cdot L = 2$  and overall cost  $T = 7$ . The effect of a reduction to  $L = 9$  is that the level of precaution is reduced to  $p = 0.9$  which implies a 50% increase in the loss probability to  $D^1 = 0.3$ , an increase in expected loss to  $D^1 \cdot L = 2.7$  and a reduction in total cost to  $T = 6.75$ .

Note that, due to re-optimization, the increase in expected utility of 0.25 exceeds the value of 0.2 that one obtains from using the naive approach of computing the value of expected quality adjusted life years using the original probability of loss. This is true in spite of a substantial "offsetting effect" (i.e., individuals reducing their level of precaution). However, since individuals in case 1 take into account all costs and benefits of precaution, the reduction in the cost of taking precaution (here,  $5 - 4.05 = 0.95$ ) is guaranteed to exceed the increase in the expected loss of 0.7. Thus, even though the offsetting effect

Case	$b$	$L$	$\alpha$	$\kappa$	$p^*$	$D$	$D \cdot L$	$c(p)$	$T$ (total cost)
case 1	1	10	5	1.2	1	0.2	2	5	7
	1	9	5	1.2	0.9	0.3	2.7	4.05	6.75
case 2	0.3	10	5	0.4	0.3	0.1	1.0	0.45	1.45
	0.3	9	5	0.4	0.27	0.13	1.17	0.365	1.535
case 3	0.98	10	10	1.4	0.49	0.91	9.1	2.4	11.50
	0.98	9	10	1.4	0.441	0.959	8.631	1.745	10.376

Table 1: The effect of improved loss mitigation technologies: Numerical examples

here is significant, the social value of the loss mitigation technology is actually larger than what the naive approach would suggest.

Of course, in cases where there is a significant negative externality, it is likely that the naive approach overestimates the social value of the loss mitigation technology. This is illustrated by case 2 in Table 1, which sets  $b = 0.3$  and  $\kappa = 0.4$ . Because  $b < 1$ , there is a negative externality effect associated with the offsetting behavior of individuals. Everyone reduces his level of precaution from 0.3 to 0.27, and this increases the probability of loss from 0.1 to 0.13, which in turn increases the expected loss from 1.0 to 1.17. Although there is a reduction in the cost of precaution, which falls from 0.45 to 0.365, the overall effect is an increase in the total cost to each individual from 1.45 to 1.535. The LMT improvement (i.e., better treatment of the disease) actually makes everyone worse off in equilibrium.

Case 3 demonstrates an interesting possibility. Here there is a small externality effect ( $b = 0.98$ ). Everyone reduces the level of precaution from 0.49 to 0.441, and that increases the probability of loss from 0.91 to 0.959. Expected loss falls from 9.1 to 8.731. This reduction of 0.369 is, of course, less than the naive expectation of 0.91 (i.e.,  $D^0(10 - 9)$ ). However, the savings in cost of precaution is 0.655 (i.e.,  $2.4 - 1.745$ ) and so the overall reduction in total cost of 1.124 ( $11.5 - 10.376$ ) exceeds the naive evaluation. This demonstrates that, even with offsetting behavior and a negative externality effect, the value of an improvement in loss mitigation technology may be larger rather than smaller than suggested by the naive view which ignores the offsetting behavior.

We now consider two examples in which an increase in  $\theta_2$  reduces the probability of loss as long as the behavior of all individuals remains constant. In our first example, we assume that an increase in  $\theta_2$  also enables individuals to be more effective when they are careful, in the sense that for any given level of  $p$ ,  $\theta$  increases the marginal benefit of being careful. Let  $D(p, \bar{p}) = [\kappa - b\theta p - (1 - b)\theta\bar{p}]L + \alpha p^2$ , so that each individual minimizes

$$[\kappa - b\theta p - (1 - b)\theta\bar{p}]L + \alpha p^2. \quad (15)$$



Solving the first order condition yields  $p = \frac{b\theta L}{2\alpha}$ , and, of course, this is equal to the equilibrium level of care by everybody else. Note that an increase in  $\theta$  here increases the equilibrium level of care. Hence, an increase of  $\theta$  has two positive effects on the individual: First, a direct one that comes through the higher effectiveness of the individual's own  $p$ . Second, this direct effect is reinforced by a positive externality effect, since all other people increase their level of care and so also reduce the loss probability for everybody else.

Formally, substituting the equilibrium level of care into the objective function yields

$$T = \left[ \kappa - \theta \frac{b\theta L}{2\alpha} \right] L + \alpha \left( \frac{b\theta L}{2\alpha} \right)^2. \quad (16)$$

Differentiating  $T$  with respect to  $\theta$  and noting that  $0 < b < 1$  yields

$$\frac{\partial T}{\partial \theta} = \frac{bL^2}{\alpha} \left( \frac{b}{2} - 1 \right) \theta < 0. \quad (17)$$

More generally, an increase in  $\theta$  is always beneficial for the individual if  $\frac{\partial D}{\partial p \partial \theta} < 0$ , i.e. if an increase in  $\theta$  increases the marginal effectiveness of being careful (i.e., from Proposition 1).

In contrast, the next example shows that an increase in  $\theta$  can be counterproductive if it makes people less careful in equilibrium. Assume now that  $D(p, \bar{p}) = (1 - \theta) [\kappa - bp - (1 - b)\bar{p}]$ . An interpretation of this function is an immunization or a treatment that is effective only for a proportion  $\theta$  of the population and prevents or heals the illness for them, while the immunization (or treatment) has no effect on the remainder of the population. Furthermore, individuals are not sure whether they personally belong to the group that is positively affected by the treatment. With this loss probability function, each individual minimizes

$$(1 - \theta) [\kappa - bp - (1 - b)\bar{p}] L + \alpha p^2. \quad (18)$$

This yields  $p = \frac{b(1-\theta)L}{2\alpha}$ , which also equals the equilibrium level  $\bar{p}$ . We assume  $0 < \theta < 1$ . Furthermore, for this example to have an interior solution, we have to assume that  $\kappa \geq \frac{b(1-\theta)L}{2\alpha}$  (to ensure that the equilibrium loss probability is non-negative). Substituting the equilibrium values of  $p$  and  $\bar{p}$  in the objective function and simplifying yields

$$T = \kappa(1 - \theta)L - \frac{b(1 - \theta)^2 L^2}{4\alpha} (2 - b). \quad (19)$$

We can differentiate this with respect to  $\theta$  to yield

$$\frac{\partial T}{\partial \theta} = L \left[ \frac{b(1 - \theta)L}{2\alpha} (2 - b) - \kappa \right]. \quad (20)$$

Using the restriction on  $\kappa$  from above, one can see that this expression may be positive if  $b$  is smaller than 1, indicating that the total cost for individuals may go up with an

increase in  $\theta$ . If, instead,  $b = 1$ , then our assumption on  $\kappa$  guarantees that  $\frac{\partial T}{\partial \theta_2} < 0$ , which implies that an increase in  $\theta$  is beneficial for individuals.

The interpretation is similar to that in the first example. An increase in  $\theta$  is directly beneficial for individuals as it reduces their loss probability and also allows them to reduce their cost of care. However, since everybody reduces his level of care, the extent of the negative externality increases and may more than offset the positive effect. The only case in which this cannot happen is if  $b = 1$ , as then an individual's loss probability depends only on his own actions.

The results derived in this subsection provide insight into how one should determine the social value of publicly provided improvements in road safety as captured by items that reduce the probability of accidents occurring (e.g., rumble strips along the edge of the road) and those that reduce the severity of accidents (e.g., improved crash barriers).

In particular, we saw that a reduction in  $L$  (i.e., a LMT improvement) leads to a reduction in self-protection. This in itself does not imply a reduction in welfare since each individual responds optimally to this improved environment and so the off-setting effect in itself does not have a detrimental welfare effect. However, the individual does not take into account the impact this reduced level of self-protection has on others' well-being through the increased accident probability for others. Thus, if this increased externality effect is strong enough, welfare will actually be reduced by such a safety improvement.

We also saw that the welfare effect of a publicly provided improvement in safety that reduces the probability of an accident (a PRT improvement) depends on whether the feature provided increases or decreases the marginal productivity of drivers' own self-protection efforts (e.g., attention paid to potential accident hazards as they arrive, speed, etc.). Since rumble strips alert drivers should they wander off the road onto the shoulder, this feature reduces the value of maintaining a high level of alertness or may induce drivers to take less frequent rest breaks. In other words, this would presumably translate into a reduction in the marginal productivity of alertness and encourage drivers to reduce their level of care. The externality effect would then reduce the value of such a *safety improvement* and we have shown that if this effect is strong enough then welfare may be reduced even if the feature is costless. On the other hand, safety improvements that enhance the marginal productivity of drivers own self-protection efforts have an added advantage through this external effect.

The welfare analysis in this subsection has demonstrated that under certain conditions the presence of offsetting behavior may lead to a reduction in social welfare that would not be recognized if behavioral effects were not taken into account. We emphasize that this general recognition is not a new finding but rather that our contribution is to point out how the functional form of  $D(p_i, \bar{p}, \theta)$  impacts on this result and how the externality

effect is crucial in understanding the overall result.

## 4.2 Endogenous Safety Improvements

The model of the previous subsection can easily be adjusted to analyze the welfare implications of safety features that can be purchased by individuals on a voluntary basis. In particular, we now investigate whether such individual choices will be socially optimal and, if not, whether government should either tax or subsidize (or prohibit or make mandatory) a given type of safety feature.

We continue to represent a person's own choice of precaution by  $p_i$  and use  $\bar{p}$  to represent the average level of precaution for the population. The variable  $\theta \geq 0$  represents the level of safety technology that affects the probability of an accident (i.e., a PRT), with  $p_i$  and  $\bar{p}$  fixed. An increase in  $\theta$  is equivalent to an improved PRT (probability reduction technology). Therefore, again we write the probability of an accident as  $D(p_i, \bar{p}, \theta)$  with  $D_j < 0$ ,  $j = 1, 2, 3$ .<sup>6</sup> As above, the variable  $\lambda \geq 0$  represents the level of safety technology that affects the magnitude of a loss,  $L(\lambda)$ , where  $L_\lambda < 0$ . An increase in  $\lambda$  is equivalent to an improved LMT (loss mitigation technology). Individuals can choose what level of PRT,  $\theta$ , or LMT,  $\lambda$ , to purchase. The cost of a given pair  $(\lambda, \theta)$  is represented by  $K(\lambda, \theta)$ ,  $K_\lambda > 0$ ,  $K_\theta > 0$ . The function  $K(\lambda, \theta)$  may in many instances be additively separable. Some innovations, however, such as an improved braking system, may affect both the probability of an accident and the size of any loss should an accident occur, and so we use the more general specification  $K(\lambda, \theta)$ . Examples of the degree of LMT ( $\lambda$ ) that an individual can choose would be the number and quality of the system of airbags, active head restraints, and quality of crash cage while examples of degree of PRT ( $\theta$ ) would be quality of brake system, rear camera, lane warning detection system.

So, first consider the individual's optimal choice of  $\lambda$ . This choice reflects the level of safety equipment that reduces severity of loss should an accident occur. The presence, number, and quality of airbags is a good example. For now we will assume  $\theta$  is fixed,<sup>7</sup> and so the individual chooses  $\{\lambda, p_i\}$  to minimize

$$\Omega(\lambda, p_i, \theta) = D(p_i, \bar{p}, \theta)L(\lambda) + c(p_i) + K(\lambda, \theta) \quad (21)$$

leading to first order conditions

$$\Omega_\lambda = D(p_i^*, \bar{p}, \theta)L_\lambda(\lambda^*) + K_\lambda(\lambda^*) = 0 \quad (22)$$

---

<sup>6</sup>As in the previous subsection, the sign of the cross-partial  $\frac{\partial^2 D}{\partial \theta \partial p_i}$  will turn out to be important and we allow for any possibility.

<sup>7</sup>Some safety features of automobiles could affect both the likelihood of sustaining an accident and the severity of the consequences. Improved brake systems may be an example. For ease of analysis we artificially maintain a distinct separation between these two types of safety features.

$$\Omega_{p_i} = D_1(p_i^*, \bar{p}, \theta)L(\lambda^*) + c'(p_i^*) = 0 \quad (23)$$

where  $\{\lambda^*, p_i^*\}$  reflect privately optimal choices. We assume  $L_\lambda < 0$ ,  $L_{\lambda\lambda} > 0$ ,  $K_\lambda > 0$ ,  $K_{\lambda\lambda}(\lambda^*) > 0$ , and further that  $\Omega$  is a convex function in the choice variables.

Equations (22) and (23) have simple interpretations. Letting  $MB^x$  and  $MC^x$  represent the private marginal benefit and cost functions for variables  $x = p_i, \lambda$  (and later  $\theta$ ), we have that

$$MB^\lambda = -D(p_i^*, \bar{p}, \theta)L_\lambda(\lambda^*) = K_\lambda(\lambda^*) = MC^\lambda \quad (24)$$

$$MB^{p_i} = -D_1(p_i^*, \bar{p}, \theta)L(\lambda^*) = c'(p_i^*) = MC^{p_i} \quad (25)$$

Note that the private benefits of course do not include any externality effects. Upon continuing the presumption that the social planner cannot influence  $p$  directly as discussed above, we want to determine whether social welfare is maximized (in a second-best sense) by the private decisions described above. In particular, will the individual choose too low or too high a level of  $\lambda$ . This is a useful question since  $\lambda$  reflects a market choice and so the government can influence this choice through a tax or subsidy.

Social welfare is given by

$$\Psi(\lambda, \theta) = -D(p(\lambda, \theta), p(\lambda, \theta), \theta)L(\lambda) - c(p(\lambda, \theta)) - K(\lambda, \theta) \quad (26)$$

where again  $\theta$  is assumed to be held fixed for now. The planner or government has no direct control over  $p$  but recognizes that the individuals' optimal choice for  $\lambda$  will influence  $p$ . Individuals ignore the external benefit of increasing  $p_i$  and this effect on welfare is reflected by the term  $D_2 < 0$ . Moreover, we have earlier shown that  $\frac{\partial p_i}{\partial \lambda} < 0$ . Thus, it is intuitive that at the privately optimal choices  $\{\lambda^*, p_i^*\}$  social welfare will rise if  $\lambda^*$  were to fall; that is, in this second-best environment individuals spend excessively on the LMT (i.e., self-insurance). To show this formally, differentiate (26) with respect to  $\lambda$ , which yields

$$\Psi_\lambda = -D(\cdot, \cdot, \cdot)L_\lambda - (D_1 + D_2)p_\lambda \cdot L - c' \cdot p_\lambda - K_\lambda. \quad (27)$$

From the first-order condition (22), we get that  $(-D_1 - c')p_\lambda = 0$  and from (23), we have that  $-D(\cdot, \cdot, \cdot)L_\lambda - K_\lambda = 0$ . Thus, evaluating (27) at the privately optimal choices  $\{\lambda^*, p_i^*\}$ , this leaves us with

$$\Psi_\lambda = -D_2 \cdot p_\lambda \cdot L < 0 \quad (28)$$

Thus, a decrease in  $\lambda$  from the privately optimal value increases social welfare.

To understand more clearly why this is so, note that in the absence of explicit control over choice of  $p_i$  it follows that the net benefit of a marginal change in  $\lambda$  can be expressed as the marginal social benefit ( $MSB^\lambda$ ) less the marginal social cost ( $MSC^\lambda$ ). Private and social benefits are the same, but the marginal social cost of  $\lambda$  includes the external cost

( $MEC^\lambda$ ) associated with the fact that an increase in  $\lambda$  leads to a reduction in care that has both a private cost and an external cost on other road users. Thus, we can write

$$MSC^\lambda = MC^\lambda + MEC^\lambda \text{ where } MEC^\lambda = D_2 \cdot p_\lambda \cdot L > 0 \quad (29)$$

From this intuition, it is clear that a standard Pigouvian tax that reflects this external cost of  $\lambda$  would be an optimal second-best intervention. This is a second-best argument because it follows only because government can't directly control  $p_i$ . If government could ensure somehow that individuals followed the first-best level in choice of driving habits, then they should be allowed to choose their privately desired level of  $\lambda$ . In the second-best setting, however, the optimal tax can be seen to be higher the greater is the response of changes in  $\lambda$  on individuals choice of self-protection as measured by  $p_\lambda$ , the greater is the external effect as measured by  $D_2$ . Although it may seem somewhat perverse, the optimal tax is proportional to  $L$ ; that is, the bigger the loss to the individual, the more any effort to reduce this loss should be taxed.

The above analysis also questions the common practice of government implementation of mandatory safety requirements when the impact of these requirements has an offsetting behavioral effect and in our model this is exaggerated by the external effect of reduced level of safe driving habits. Our results here represent explicit treatment of concerns implicitly treated by, for example, Peltzman (1975) and Boyer and Dionne (1987).<sup>8</sup>

Now consider a similar exercise but involving an individual's optimal choice of  $\theta$ . This choice reflects the level of some safety equipment or feature of an automobile that reduces the probability of an accident but has no effect on severity of loss. The quality of brakes or presence and quality of anti-skid or stability system are possible examples. Here we assume  $\lambda$  is fixed (and sometimes subsumed in notation) and the individual chooses  $\{\theta, p_i\}$  to minimize  $\Omega(\lambda, p_i, \theta)$  given in equation (21), leading to first-order conditions

$$\Omega_\theta = D_3(p_i^*, \bar{p}, \theta^*)L(\lambda) + K_\theta(\theta^*) = 0 \quad (30)$$

$$\Omega_{p_i} = D_1(p_i^*, \bar{p}, \theta^*)L(\lambda) + c'(p_i^*) = 0 \quad (31)$$

where  $\{\theta^*, p_i^*\}$  reflect privately optimal choices. We assume  $D_3 < 0$ ,  $D_{33} > 0$ ,  $K_\theta > 0$ ,  $K_{\theta\theta}(\lambda^*) > 0$ , and further that  $\Omega$  is a convex function in the choice variables.

As before, the first-order conditions have simple interpretations involving marginal benefit and cost of each variable  $p_i$  and  $\theta$ .

$$MB^\theta = -D_3(p_i^*, \bar{p}, \theta^*)L(\lambda) = K_\theta(\theta^*) = MC^\theta \quad (32)$$

---

<sup>8</sup>Boyer and Dionne (1987) point out that policies such as compulsory wearing of seatbelts may indeed lead to less careful driving. Although they hold the size or severity of loss constant, they do consider the effect of such a policy on  $p_i$  in a reduced form manner consistent with our equilibrium function  $p(\lambda, \theta)$  with  $p_\lambda < 0$ . They consider a broader range of policies to influence the individual's choice of  $p_i$  than we do.

$$MB^{p_i} = -D_1(p_i^*, \bar{p}, \theta^*)L(\lambda^*) = c'(p_i^*) = MC^{p_i} \quad (33)$$

Note, of course, that the private benefits do not include any externality effects. Again, continuing the presumption that the social planner cannot influence  $p$  directly as discussed above, we want to determine whether social welfare is maximized (in a second-best sense) by the private decisions described above.

Differentiating social welfare  $\Psi(\lambda, \theta)$  in equation (26) with respect to  $\theta$ , and evaluating at the privately optimal choices  $\{\theta^*, p_i^*\}$ , we have

$$\Psi_\theta = -(D_1 + D_2) \cdot p_\theta \cdot L - D_3 \cdot L - c' \cdot p_\theta - K_\theta. \quad (34)$$

This can be rewritten as

$$\Psi_\theta = -(D_1 \cdot L + c') \cdot p_\theta - (D_3 \cdot L + K_\theta) - p_\theta \cdot D_2 \cdot L \quad (35)$$

Upon collecting terms and evaluating this derivative at the privately optimal choices  $\{\lambda^*, p_i^*\}$ , we can see from the first-order conditions above (w.r.t.  $p_i$  and  $\theta$ ) that the first two terms (in brackets) are zero. This leaves us with

$$\Psi_\theta = -p_\theta \cdot D_2 \cdot L \quad (36)$$

The sign of  $\Psi_\theta$  is the same as the sign of  $p_\theta$ . Thus, use of  $\theta$  should be taxed if it reduces marginal effectiveness of  $p$  (safe driving habits) but should be subsidized if it increases marginal effectiveness of  $p$ .

While this may seem intuitively clear, one should remember that it is not the effect of  $\theta$  on the marginal impact of safer driving habits per se that is important but only through its relationship with the external effect of  $p$ . That is, if  $D_2 = 0$  then the privately optimal choice for  $\theta$  would also be the socially optimal choice regardless of whether  $\theta$  improves or worsens the marginal effectiveness of safe driving habits. In other words, if there are no externalities on other individuals, then the existence of some off-setting behavior is of no relevance to the social planner's problem.

As with the choice of  $\lambda$ , the above relationships can be expressed in terms of the marginal social benefit ( $MSB^\theta$ ) and the marginal social cost ( $MSC^\theta$ ) of  $\theta$ . Private and social benefits can be equated but marginal social cost of  $\theta$  includes the external cost, or in this case possibly benefit, ( $MEC^\theta$ ) associated with the fact that an increase in  $\theta$  may lead to either a reduction or increase in care that has both a private cost and an external cost on *other* road users. Thus, we can write

$$MSC^\theta = MC^\theta + MEC^\theta \text{ where } MEC^\theta = D_2 \cdot p_\theta \cdot L > 0. \quad (37)$$

To summarize, individual choices involving expenditure on automobile safety features that reduce the size of loss (self-insurance) are excessive when driving habits cannot be

observed or otherwise controlled. Thus, a tax on the cost of any such feature should be applied.

Individual choices involving expenditure of safety features that reduce the probability of getting into a traffic accident should be taxed only in the case where these features reduce the marginal productivity of unobserved self-protective activities (i.e., safer driving habits). If such a feature has a neutral effect on marginal productivity of safer driving habits, then an individual's optimal choice will also be socially optimal. If such a feature increases the marginal productivity of safer driving habits, then the individual will under-consume such a feature and it should be subsidized.<sup>9</sup>

## 5 Conclusion

The implications of offsetting behaviour turn out to be important in many policy and regulatory frameworks such as traffic safety and infectious diseases. We have developed a simple unified model that allows us to analyze positive and normative implications of offsetting behavior, and offers a useful starting point for the study of many types of applications. The effects of offsetting behavior depend on (1) whether individuals' activity levels are strategic complements or substitutes; (2) how the safety technology affects the marginal value of precaution; (3) the strength of the externality effect; and (4) whether the adopted safety technology mitigates the size of loss (LMT) or reduces the probability of loss (PRT).

We are interested in the positive and normative effects of improved probability reduction or loss mitigation technologies, either imposed through regulation or chosen by the individual at some cost. We demonstrate that the size of the offsetting behaviour effect on the loss probability is not a useful indicator of whether an improved safety technology, whether of the LMT or PRT type, improves or worsens welfare. Nor is the size of the offsetting effect a useful indicator for the size of the welfare effect.

For example, the naive estimate of the welfare gain from a LMT improvement that reduces the size of the loss by  $\Delta L$  is  $D_0 \cdot \Delta L$ , where  $D_0$  is the initial loss probability without any offsetting behaviour. Even if the offsetting behaviour effect is so strong as to lead to an increased expected loss, it is possible that a welfare gain occurs and that its size exceeds this naive estimate ( $D_0 \cdot \Delta L$ ). This type of result demonstrates the importance of not only measuring the empirical size of offsetting effects but also accounting for all

---

<sup>9</sup>This analysis is similar in spirit to that of Arnott and Stiglitz (1986), although the source of inefficiency due to moral hazard in their model operates through the insurance market rather than a classical externality effect. They have a similar result, however, in that whether one should tax or subsidize a commodity that reduces the probability of an accident depends on how that commodity affects the marginal productivity of (unobserved) effort. See their Proposition 2, p. 12.

effects in the context of a properly specified model.

We also analyze a setting where different levels of technologies for LMTs and PRTs, such as quality of safety features for automobiles, are available at private cost to individuals. We show that, whether it is welfare enhancing to tax or subsidize the technology purchase depends on whether the effect of an increase in its quality enhances or reduces the marginal benefit of each individual's own level of personal effort choice (i.e., such as driving speed). Moreover, the size of the optimal tax depends on the strength of all features of the model described above.

There are many avenues for future research based on our modeling approach. For example, papers by Boyer and Dionne (1987), Gossner and Picard (2005), and Muermann and Kunreuther (2009), develop specific models of road safety in the presence of insurance and analyze how the choice of insurance schemes, such as fault versus no fault, affects welfare. While our model is set up with non-monetary (uninsurable) losses, it would certainly be interesting to analyze some of the questions and policies of those papers within the context of our model.

It would also be interesting to extend our model by introducing differences between individuals in their preferences for loss mitigation or cost of taking precaution. Doing so would help to inform empirical models that attempt to explain why individuals choose different levels of such technologies and would lead to nontrivial extensions of the normative implications of technology improvements in the presence of offsetting behaviour.

## References

- Arnott, A. and J. Stiglitz (1986): "Moral Hazard and Optimal Commodity Taxation", *Journal of Public Economics*, vol. 29, pp. 1-24.
- Blomquist, G. (1986): "A utility Maximization Model of Driver Traffic Safety Behavior", *Accident Analysis and Prevention*, vol. 18(5), pp. 371-375.
- Boyer, M. and G. Dionne (1987): "The Economics of Road Safety", *Transportation Research B*, vol. 21B, pp. 413-31.
- Cooper, R. and T. W. Ross (1985): "Product Warranties and Double Moral Hazard", *Rand Journal of Economics*, pp. 103-113.
- Ehrlich, I., and G. Becker (1972): "Market Insurance, Self-insurance, and Self-protection", *Journal of Political Economy*, vol. 80, no. 4, pp. 623-48.
- Fletcher, J. M., Frisvold, D. E., N. Tefft (2010): "The Effects of Soft Drink Taxes on Child and Adolescent Consumption and Weight Outcomes", *Journal of Public Economics*, vol. 94, pp. 967-974.
- Fortin, B. and P. Lanoie (1992): "Substitution Between Unemployment Insurance and



- Workers' Compensation", *Journal of Public Economics*, vol. 49, pp. 287-312.
- Geoffard, P.-Y. and T. Philipson (1996): "Rational Epidemics and Their Public Control", *International Economic Review*, vol. 37, no. 3, pp. 603-624.
- Gossner, O. and P. Picard (2005): "On the Consequences of Behavioral Adaptations in the Cost-Benefit Analysis of Road Safety Measures", *Journal of Risk and Insurance*, vol. 72, no. 4, pp. 557-599.
- Harless, D. W., and G. E. Hoffer (2003): "Testing for Offsetting Behavior and Adverse Recruitment Among Drivers of Airbag-Equipped Vehicles", *Journal of Risk and Insurance*, vol. 70, no. 4, pp. 629-650.
- Hause, J. C. (2006): "Offsetting Behavior and the Benefits of Safety Regulations", *Economic Inquiry*, vol. 44(4), pp. 689-698.
- Holmstrom, B. (1982): "Moral Hazard in Teams", *Bell Journal of Economics*, vol. 13, pp. 324-340.
- Kunreuther, H. and G. Heal (2003): "Interdependent Security", *The Journal of Risk and Uncertainty*, vol. 26(2/3), pp. 231-249.
- Lanoie, P. (1992): "The Impact of Occupational Saafety and Health Regulation on the Risk of Workplace Accidents: Quebec, 1983-87", *The Journal of Human Resources*, vol. 27, no. 4, pp. 643-660.
- McCannon, B. G. (2010): "Strategic Offsetting Behavior: Evidence from National Collegiate Athletic Association Men's Basketball", *Contemporary Economic Policy*, vol. 29(4), pp. 530-563.
- Miljkovic, D., Nganje, W., and B. Onyango (2009): "Offsetting Behaviour and teh Benefits of Food Safety Regulation", *Journal of Food Safety*, pp. 49-58.
- Muermann and Kunreuther (2008): "Self-protection and Insurance with Interdependencies", *Journal of Risk and Insurance*, vol. 36 pp. 103-123.
- Neill, J. R. (1993): "A Theoretical Reappraisal of the Offsetting Behavior Hypothesis", *Journal of Regulatory Economics*, vol. 5, pp. 435-440.
- P. A. Pedersen (2003): "Moral Hazard in Traffic Games", *Journal of Transport Economics and Policy*, vol. 37, pt.1, pp. 47-68.
- S. Peltzman (1975): "The Effects of Automobile Safety Regulation", *Journal of Political Economy*, vol. 83, no. 4, pp. 677-725.
- T. Philipson (2000): "Economic Epidemiology and Infectious Diseases", in (eds. A. J. Culyer and J. P. Newhouse) *Handbook of Health Economics*, vol. 1B, North Holland, Elsevier, Amsterdam, pp. 1761-1799.
- Potter (2011): "Estimating the Offsetting Effects of Driver Behavior in Response to Safety Regulation: The Case of Formula One Racing", *Journal of Quantitative Analysis in Sports*, vol. 7(3), pp. 1-20.

Risa, A. (1992): “Public Regulation of Private Accident Risk: The Moral hazard of Technological Improvements”, *Journal of Regulatory Economics*, vol. 4, no. 4, pp. 335-346.

A. E. Risa (1995): “The Welfare State as Provider of Accident Insurance in the Workplace: Efficiency and Distribution in Equilibrium”, *The Economic Journal*, vol. 105, no. 428, pp. 129-44.

Winston, C. V. Maheshri, and F. Mannering (2006): “An Exploration of the Offset Hypothesis Using Disaggregate Data: The Case of Airbags and Antilock Brakes”, *Journal of Risk and Uncertainty*, vol. 32, pp. 83-99.