2015

The value of technology improvements in games with externalities: A fresh look at offsetting behavior

Michael Hoy, *University of Guelph*
Mattias K Polborn

The value of technology improvements in games with externalities: A fresh look at offsetting behavior

Michael Hoy*
Department of Economics and Finance
University of Guelph
Guelph, ON N1G 2W1, Canada
e-mail: mhoy@uoguelph.ca

Mattias K. Polborn
Department of Economics
University of Illinois
216 David Kinley Hall, 1407 W. Gregory Dr.
Urbana, IL 61801, USA
e-mail: polborn@uiuc.edu

June 24, 2015

Abstract

We model the effect of safety technology improvements in a symmetric game in which each player's payoff depends on his own precaution and the other players average precaution. We derive conditions under which an improved technology increases or decreases players' equilibrium utility.

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 technology improvement is a complement or substitute for individual precaution. For safety technologies that individuals can choose whether or not to purchase, individuals expend too much on reducing the loss size but may spend either too much or too little on features that reduce the individual's loss probability.

JEL Codes: C7, D6, D7, D8

Keywords: offsetting behavior, moral hazard

*Corresponding author: phone (+1)519-763-0221

1 Introduction

The seminal contribution of Peltzman (1975) addresses the possibility that legally mandated use of safety devices on automobiles such as seat belts may lead to offsetting effects in the form of reduced care in driving habits. Such an individually rational response to the lowered cost or severity of accidents can partially or even entirely offset the anticipated reductions in the overall accident and/or fatality rate. This type of phenomenon is known as offsetting behavior or the risk compensation effect.

Following Peltzman (1975), numerous studies analyze whether adoption of new safety technologies leads to offsetting behavior in the context of road safety (e.g., Winston et al. (2006) and Harless and Hoffer (2003)) as well as many other areas such as 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.), health (e.g., Geoffard and Philipson (1996), Fletcher et al. (2010)).

In most real world applications of offsetting behavior, an important externality is present. A driver's accident probability, for example, depends crucially on how carefully others drive, in addition to the individual's own level of care and use of technology. The likelihood of incurring an infectious disease also depends on how careful others are (or have been) in avoiding the disease since this affects the percentage of the population that is infected. However, in many empirical analyses of offsetting behavior, the externality is typically either not explicitly modeled or even recognized. In others, the externality is described in a way that suits only the particular application under study.

In order to evaluate the welfare consequences of safety technology innovations when externalities are present, we need to understand how each individual's equilibrium choice of precaution reacts to the technological innovation, but also how his utility is affected by the choices made by others (which are also affected by the same technological innovation).

This is the purpose of this paper. In our model, we assume that the probability of an individual experiencing a loss depends both on the individual's own level of precaution and, either positively or negatively, on the average precaution level of other agents.[1] An improved safety technology generally affects agents' equilibrium level of care – either positively or negatively, depending on whether the innovation is a substitute or complement to precaution. The welfare effect is then the sum of the direct effect of the innovation, which would arise if an individual was the only one with access to the new technology, and

[1] By an individual's level of precaution, we mean things such as attentiveness to road hazards while driving or use of safe sex practices. These are assumed unobservable to the social planner, and create externalities for others. 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 (1992), Pedersen (2003), and Risa (1992, 1995) for specific applications.

the indirect effect that works through changing other individuals' equilibrium precaution level.

Per se, offsetting behavior does not pose a normative problem. In fact, in the absence of an externality effect, offsetting behavior actually contributes positively to welfare and simply reflects people re-optimizing due to the change in the safety environment. However, if everyone's loss probability is decreasing in other agents' equilibrium level of care, and if the new safety technology reduces the equilibrium level of care, then this effect is detrimental and needs to be accounted for in welfare comparisons. Indeed, it is even possible that the overall welfare effect of a technology improvement is negative.

We also develop two specialized versions of our model that help us to relate our approach to the existing literature. The first one of these is what we call a *loss mitigation technology* (LMT). It reduces the size, but not the probability, of a loss (think of seat belts or airbags). The second one, a *probability reduction technology* (PRT), reduces the probability of a loss, but not its size if it occurs (think of rumble strips on highways that warn a driver if he is about to leave his lane and sometimes allow for corrective actions).

We show that an improved LMT always leads to a reduction in the equilibrium level of precaution, consistent with Peltzman's hypothesis. If the externality from precaution is a positive one, as is plausible in the context of traffic safety, then the offsetting behavior has negative welfare consequences. However, in applications where the externality is a negative one, such as in certain types of crime deterrence, then the offsetting behavior actually has a positive welfare effect.

Whether an improved PRT leads to offsetting behavior in the traditional sense (i.e., a reduction in individuals' precaution) or the opposite depends on how the technology affects the marginal effectiveness of precaution and can go in either direction. For example, an improved braking system may increase the marginal benefit of attentiveness for avoiding accidents which could not otherwise be avoided,[2] while rumble strips may reduce the individual's perceived value of more frequent rests while driving. The welfare effect of a PRT then depends on both the direction of the effect on precaution and on whether the externality of precaution is positive or negative.

In Section 2, we develop our unified model for analyzing the positive and normative implications of offsetting behavior. We consider both cases in which the level of safety technology is exogenously imposed or endogenously chosen. In Section 3, we address the issue of valuing discrete, exogenous changes in safety technology and compare it to the naive or engineering approach that ignores behavioral effects. In the context of specific models that reflect LMTs and PRTs, we analyze when the naive approach overestimates or underestimates the true value. Section 4 provides a discussion of our results, including

[2]Of course, an improved braking system could also induce people to drive faster or less carefully.

how our model relates to the existing literature. Section 5 concludes and offers suggestions for further research on this topic.

2 The unified model

2.1 Setup

We now present a model that unifies many existing models of offsetting behavior. Our model is first developed for the case of exogenous (mandatory) safety technologies, such as improved crash barriers on roadways or mandatory seat belt legislation. However, we also extend the model to allow for safety technologies that are chosen and paid for by individuals, such as airbag systems.

Consider a game between a continuum of players. All players are symmetric and each player's payoff depends on his own activity level x (which we refer to as his level of precaution), the average activity level $\bar{x}$ of the other players in equilibrium, and technology θ. Each player's problem therefore is

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

We assume that f is strictly concave in the first argument ($f_{11} < 0$)[3], and increasing in its third argument ($f_3 > 0$): 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.

In many activities, such as automobile driving, the externality is plausibly positive, i.e., any individual's probability of incurring a bad outcome is reduced by others taking more care. However, a negative externality with respect to precaution is also a possibility. For example, an individual's own probability of being burglarized may increase (and hence his welfare decrease) as a result of others increasing their level of observable precaution. This effect may follow because when others make their houses less attractive to burglars (e.g., by ensuring house lights are automatically turned on and off while away), burglars may shift their attention to seemingly more vulnerable properties.[4]

Note that our model is set up with uninsurable losses in mind; for example, losses could reflect lost quality adjusted life years. Furthermore, our model implicitly assumes that the individual choice of precaution cannot directly be controlled by the social planner (because this is either impossible or very costly to effect), and the level of indirect measures such

[3]Henceforth, we will denote partial derivatives by subscripts. For example, $f_{11} \equiv \partial f^2/\partial x^2$.

[4]Whether a potential victim's precaution creates a positive or negative externality depends on whether the action is observable or unobservable to perpetrators of crime. See Ayres and Levitt (1998) and Shavell (1991).

as experience rating by insurers, liability through negligence rules enforced through the legal system, or imperfect monitoring such as police enforcement of traffic regulations[5] is constant with respect to the safety technology improvement. Our reason is that we wish to analyze the effect of technological change on the externality created by moral hazard and on social welfare in isolation of other issues. We recognize that, even without this direct type of externality, individual moral hazard can create a negative externality effect through an insurance pool (e.g., see Gossner and Picard, 2005). We leave aside these sorts of issues in this article, although they are all well worth exploring in future work.

2.2 Analysis

Since f is strictly concave in its first argument, the first order condition

$$\frac{\partial f}{\partial x} \equiv f_1 = 0 \tag{2}$$

is necessary and sufficient for a global optimum. An important property of the equilibrium is stability for constant technology, i.e., whether, when the other players' average action increases by 1, the individually optimal action increases by less than 1. Applying the implicit function theorem to (2) to calculate the slope of the reaction function, we assume that

$$\frac{dx}{d\bar{x}} = -\frac{f_{12}}{f_{11}} < 1 \tag{3}$$

in order to guarantee stability. Note that this includes three qualitatively different cases: In the first case, precaution by others reduces the marginal effect of individual precaution ($f_{12} < 0$ is sufficient for (3) to hold, given that $f_{11} < 0$). In the second case, $0 < f_{12} < -f_{11}$, precaution by others increases the marginal effect of individual precaution. Finally, if $f_{12} = 0$, the other players' precaution level does not influence the marginal effect.

Totally differentiating (2) with respect to θ, we find that

$$\frac{dx}{d\theta} = -\frac{f_{13}}{f_{11} + f_{12}}. \tag{4}$$

Given (3), the denominator in (4) is negative, so the sign of $dx/d\theta$ is the same as the sign of f_{13}. Intuitively, if an increase in θ reduces the marginal payoff to x (i.e., if $f_{13} < 0$), then the equilibrium level of x falls and vice versa.

Empirical investigations of offsetting behavior usually presume that $\frac{dx}{d\theta} < 0$. However, "negative offsetting behavior," i.e. a case in which the improvement in safety is reinforced by an increase in precaution so that $\frac{dx}{d\theta} > 0$, is certainly a possibility.

We now turn to the effect of a technological improvement on the players' equilibrium utility,

$$V(\theta) = f(\bar{x}(\theta), \bar{x}(\theta), \theta). \tag{5}$$

[5] See, for example, Boyer and Dionne (1987) for an exploration of the some of these measures.

Using the envelope theorem, a change of θ has the following effect on equilibrium utility:

$$V'(\theta) = \frac{\partial f}{\partial \theta} + \frac{\partial f}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial \theta} = f_3 - f_2 \frac{f_{13}}{f_{11} + f_{12}}. \tag{6}$$

Remember that $f_3 > 0$; that is, increasing θ means progress: Holding constant the actions of all players, each player's payoff increases. The second term measures the effect of interaction between players. If the externality between players is positive ($f_2 > 0$), then the interaction effect has the same sign as 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. 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 increases the negative externality and 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. Proposition 1 summarizes these findings.

Proposition 1. *The following are sufficient conditions for a costless and exogenous (mandatory) safety improvement to increase the players' equilibrium utility level:*

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 ($f_2 = 0$).*

If none of the above conditions holds, then the equilibrium utility change is smaller than the direct effect of technical progress. In particular, the overall effect is negative if

$$f_3 < f_2 \frac{f_{13}}{f_{11} + f_{12}}. \tag{7}$$

Consider now the case that the safety technology is endogenous, i.e., chosen and paid for by the individual. Individuals internalize the financial cost of acquiring any particular level of safety technology, as well as their own benefit from the better technology, but not any externality, should one exist. All the partial derivatives above reflect the same considerations and interpretations. However, given that θ is chosen by the individual at a cost $k(\theta)$, our original problem (equation (1)) becomes $\max_{\{x,\theta\}} f(x, \bar{x}, \theta) - k(\theta)$ and the following additional first-order condition arises;

$$f_3 - k'(\theta) = 0. \tag{8}$$

The most relevant welfare question is whether a tax or subsidy on the safety technology is optimal. Given that the technology is costly, there is an additional term $-k'(\theta)$ in (6), and we have to take into account (8), which leads to:

$$V'(\theta) = \frac{\partial f}{\partial \theta} - k'(\theta) + \frac{\partial f}{\partial \bar{x}} \frac{\partial \bar{x}}{\partial \theta} = -f_2 \frac{f_{13}}{f_{11} + f_{12}}. \tag{9}$$

The term in (9), $-f_2 \frac{f_{13}}{f_{11}+f_{12}}$, captures the marginal externality effect of the safety technology on the marginal value of the individual's precaution. With no externality effect ($f_2 = 0$), and/or no effect of the safety technology on the marginal value of the individual's precaution ($f_{13} = 0$), the privately optimal choice is also the socially optimal choice, and no tax or subsidy on θ is warranted.

If f_2 and f_{13} have the same sign, then the marginal welfare effect of an increase in θ, evaluated at the private optimum, is positive. Consequently, subsidizing the safety technology is welfare-improving. Conversely, if f_2 and f_{13} have different signs, then the marginal welfare effect of an increase in θ is negative, and a tax on the safety technology is welfare-improving. These results are summarized in Proposition 2.

Proposition 2. *Suppose individuals choose privately a level of safety technology, at a cost of $k(\theta)$. If either $f_2 = 0$, or $f_{13} = 0$, or both, hold, then the privately optimal choice is also the socially optimal choice. Otherwise, we have the following implications regarding a tax or subsidy on θ being welfare improving.*

1. *If the externality is positive ($f_2 > 0$) and technical progress decreases the equilibrium activity level ($f_{13} < 0$), then $V'(\theta) < 0$ and so an appropriate tax on θ will increase welfare. The same applies if $f_2 < 0$ and $f_{13} > 0$.*

2. *If the externality is positive ($f_2 > 0$) and technical progress increases the equilibrium activity level ($f_{13} > 0$), then $V'(\theta) > 0$ and so an appropriate subsidy on θ will increase welfare. The same applies if f_2 and f_{13} are both negative.*

In any empirical analysis of offsetting behavior, the objective is generally to determine if $\frac{dx}{d\theta}$ is nonzero and, if so, estimating its magnitude. Propositions 1 and 2 imply a number of important consequences relating to both positive and normative implications of improved safety technologies.

Firstly, suppose that there is no externality effect associated with individuals' choice of precaution ($f_2 = 0$). If $f_{13} \neq 0$, there is an offsetting effect, but it is of no consequence to social welfare, irrespective of the magnitude of $\frac{dx}{d\theta}$. This follows independently of whether the improved safety technology is being mandated or chosen voluntarily by individuals.

Secondly, if the improved safety technology does not affect the equilibrium level of precaution (i.e., $f_{13} = 0$), then there also is no welfare issue, even if there is an externality associated with precaution (i.e., $f_2 \neq 0$).[6]

Propositions 1 and 2 imply that what is critically important in assessing the normative properties of the safety improvement is the interaction of its effect on the marginal benefit of precaution (f_{13}) with the externality effect (f_2). Using (4), the normative impact in both propositions (see equations (6) and (9)) can be written simply as

$$f_2 \frac{dx}{d\theta} = -f_2 \frac{f_{13}}{f_{11} + f_{12}}. \tag{10}$$

The sign and size of (10) clearly depends crucially on the sign and size of the externality effect f_2. For the sake of clarity, we now focus on the perhaps more typical case that other people taking more care increases an individual's utility ($f_2 > 0$) and a truly offsetting effect ($f_{13} < 0$, so that $\frac{dx}{d\theta} < 0$).[7]

The sign and magnitude of the cross partial f_{12} influences the size of the observable offsetting effect ($\frac{dx}{d\theta}$) as follows. The size of $\frac{dx}{d\theta}$ is inversely related to the magnitude of ($f_{11} + f_{12}$). Recall that $f_{11} < 0$ by the assumption that the objective function is strictly concave in an individual's level of precaution. Therefore, our stability condition (3) is always satisfied for $f_{12} \leq 0$ and allows for $f_{12} > 0$ but "not too large."

If precaution is a strategic substitute ($f_{12} < 0$), this leads to an amelioration of this negative welfare effect: In this case, each individual responds by increasing his own level of precaution (x) when others become less careful ($\overline{x} \downarrow$). Thus, although each person reduces his level of precaution due to its marginal benefit falling ($f_{13} < 0$), this effect is diminished by the response to others' reduced $\bar{x}$. The more negative is f_{12}, the greater is the magnitude of the denominator of $\frac{dx}{d\theta}$ and so the smaller is the offsetting effect. This also means that, for a given value of f_2, welfare is affected less.

If precaution is a strategic complement ($f_{12} > 0$), a stronger complementarity leads to a larger magnitude of the offsetting effect and a more negative welfare effect of the offsetting behavior. Intuitively, the improved safety technology leads directly to a lower equilibrium precaution level, and this decrease is exacerbated by the strategic complementarity, leading to more drastic negative welfare consequences. Formally, a larger value of f_{12} decreases the absolute value of the denominator in (10) and thus increases the size of the external effect.

Finally, if $f_{12} = 0$ then the decisions of others' affect each individual's own safety, but the individual marginal benefit of precaution is unaffected by the decisions of others. Consequently, the welfare effect is intermediate between the two cases of strategic substitutability and strategic complementarity.

[6] Note that an empirical analysis that finds no offsetting effect essentially uncovers an instance of $f_{13} = 0$ and hence does not imply any misleading welfare conclusions about offsetting behavior.

[7] It is straightforward to adjust the arguments given to either or both $f_2 < 0$ and $f_{13} > 0$.

Our model described so far provides a fairly general framework for analyzing the welfare implications of safety technology improvements in the presence of an offsetting phenomenon. The most important message is that it is not sufficient per se to determine the empirical effect of a change in some safety technology on the equilibrium choice of precaution to understand its normative significance. It is also critical to examine the relationship of individuals' decisions and the sign and size of the externality effect.

3 Valuing a discrete change in safety

In this section, we are interested in determining the true value of a discrete mandatory safety improvement, and compare it to the commonly used *engineering approach* that ignores behavioral changes induced by safety improvement.

A discrete change in safety technology is modeled as an increase in θ from θ_0 to $\theta_1 > \theta_0$. Letting $x_0 = x(\theta_0)$ and $x_1 = x(\theta_1)$, it follows that the true social value of such a change is

$$TV \equiv V(\theta_1) - V(\theta_0) = f(x_1, x_1, \theta_1) - f(x_0, x_0, \theta_0). \tag{11}$$

Using the total derivative of the value function $V(\theta) = f(x(\theta), x(\theta), \theta)$, which is

$$\frac{dV}{d\theta} = (f_1 + f_2)\frac{dx}{d\theta} + f_3, \tag{12}$$

and the fact that $f_1 = 0$ by optimality, we can write (11) as

$$TV = \int_{\theta_0}^{\theta_1} \frac{dV}{d\theta} d\theta = \int_{\theta_0}^{\theta_1} \left[f_2 \frac{dx}{d\theta} + f_3 \right] d\theta. \tag{13}$$

The term f_3 represents the direct effect of the technology improvement, conditional on no change in equilibrium precaution, while the term $f_2 \frac{dx}{d\theta}$ in (13) represents the effect due to changes in the equilibrium externality level caused by the improved safety technology.

In contrast, the engineering approach calculates the naive value as

$$NV \equiv f(x_0, x_0, \theta_1) - f(x_0, x_0, \theta_0) = \int_{\theta_0}^{\theta_1} f_3(x_0, x_0, \theta) d\theta. \tag{14}$$

This value is a naive assessment in the sense that it assumes that there is no effect on behavior and so x is held fixed at the level x_0.

Just as in (13), the term f_3 represents the direct effect of the technology improvement. Note, however, that f_3 in (13) and (14) are evaluated at different values. This subtle observation is the reason why TV and NV may differ even in settings where there is no externality ($f_2 = 0$). Specifically, the naive assessment underestimates the social value in this case, as it ignores that individuals enhance their utility by reoptimizing under the new technology.

To see this, consider first the case that the technology improvement is a substitute for precaution, and thus optimal precaution decreases (i.e., $f_{13} < 0$, so that $x(\theta) < x_0$ for all $\theta > \theta_0$). Using $f_{13} < 0$, we have that $f_3(x(\theta), x(\theta), \theta) > f_3(x_0, x_0, \theta)$.[8] Hence, the integrand in (13) is larger than the integrand in (14) for all $\theta \in (\theta_0, \theta_1]$, and thus $TV > NV$.

Second, consider the case that the technology improvement is a complement for precaution, and thus optimal precaution increases (i.e., $f_{13} > 0$, so that $x(\theta) > x_0$ for all $\theta > \theta_0$). Again using $f_{13} > 0$, this implies that $f_3(x(\theta), x(\theta), \theta) > f_3(x_0, x_0, \theta)$, and so the integrand in (13) is again larger than the integrand in (14) for all $\theta \in (\theta_0, \theta_1]$, and thus $TV > NV$. Therefore, we have just shown Proposition 3.

Proposition 3. *If $f_2 = 0$, i.e., if there is no externality from precaution, then the naive engineering approach underestimates the true value of a safety innovation whenever there is an offsetting effect (i.e., whenever $f_{13} \neq 0$).*

We now turn to two specific functional forms of our general model. The first one of these is what we call a *loss mitigation technology* (LMT). It reduces the size of the loss, conditional on one occurring, but not its probability. For example, seat belts or airbags are appropriately thought of as LMTs. Another example from the field of health economics is improved HIV treatment.

The second one, a *probability reduction technology* (PRT) reduces the probability of a loss, but not its size if it occurs. One can think of rumble strips on highways that warn an inattentive driver through the noise that the tires create on them that he is about to leave his lane, which may sometimes allow for corrective actions.

Even though, in reality, most safety technology improvements affect both the probability and the size the loss simultaneously, these two cases are convenient for their analytic tractability and characterize much of the empirical research on offsetting behavior.

3.1 Loss Mitigation Technology

For a loss mitigation technology, the size of the loss is a function of the safety technology θ. Formally, we have $L(\theta)$, with $L'(\theta) < 0$.; i.e., an increase in θ represents improved loss mitigation. The probability of a loss, $D(x, \overline{x})$, is independent of the technology and depends both on the individual's precaution, and on the average level of precaution of

[8] Note that the second argument of these functions has no effect on their value because $f_2 = 0$ by assumption, and therefore also $f_{32} = 0$.

others, $\overline{x}$. Individuals incur personal cost $c(x)$ for their level of precaution, with $c'(x) > 0$.[9]

This LMT model exemplifies many that are used in the literature, either explicitly or implicitly, to analyze offsetting behavior resulting from improved loss mitigation technologies. The externality, however, is usually not included in these models.

The individual's utility is the negative of expected loss and cost of precaution, and so we have

$$f(x, \overline{x}, \theta) = -D(x, \overline{x})L(\theta) - c(x) \tag{15}$$

The first-order condition that relates to equation (2) is:

$$f_1 = -D_1(x, \overline{x})L(\theta) - c'(x) = 0 \tag{16}$$

and implies that the individual's choice of precaution equates the marginal benefit of precaution (i.e., the reduction in expected loss) to its marginal cost.

As is typical in such models, we assume increasing precaution reduces the probability of loss at a decreasing rate ($D_{11} > 0$) and that marginal cost is increasing in the level of precaution ($c'' > 0$). These assumptions ensure that the second order condition to the maximization problem is satisfied.

The value function is

$$V(\theta) = f(x(\theta), x(\theta), \theta) = -D(x(\theta), x(\theta))L(\theta) - c(x(\theta)). \tag{17}$$

Again, we let $x_0 = x(\theta_0)$ and $x_1 = x(\theta_1)$. We also use $D^i = D(x_i, x_i)$ to represent the equilibrium loss probability when $x = x_i$, $i = 0, 1$ and $L^i = L(\theta_i)$, to represent the loss for $i = 0, 1$. Therefore, $L^0 - L^1 > 0$ is the loss reduction due to the LMT while $D^1 - D^0$ is the change in the probability due to the change in precaution.

Consequently, the naive valuation is

$$NV \equiv f(x_0, x_0, \theta_1) - f(x_0, x_0, \theta_0) = D^0\left(L^0 - L^1\right) \tag{18}$$

which is simply the expected value of the reduced loss *conditional* on behavior not changing. The true valuation is

$$TV = V(\theta_1) - V(\theta_0) = D^0\left(L^0 - L^1\right) - \left(D^1 - D^0\right)L^1 + \left[c(x_0) - c(x_1)\right]. \tag{19}$$

Since an improved loss mitigation technology reduces the marginal benefit of precaution (i.e., $L'(\theta) < 0 \Rightarrow f_{13} = -D_1(x, \overline{x})L'(\theta) < 0$), we have $\frac{dx}{d\theta} < 0$. Suppose an increase

[9] If the loss is a non-monetary object such as quality adjusted life years (QALYs), then other objects in the objective function f, such as cost of precaution $c(x)$, need to be treated as having been converted to, or measured by reference to, similar units. See Hammitt (2005) and Hirth et al. (2000) for useful discussions on linking QALYs to monetary values.

in others' precaution also reduces the probability of loss ($D_2 < 0$). This means that others' precaution represents a positive externality ($f_2 > 0$), and is a sufficient condition for $D^1 > D^0$; that is, the reduction in loss size leads to an increased loss probability due to offsetting behavior.

Under these assumptions, the true valuation equals the naive valuation, $D^0 \left(L^0 - L^1\right)$, plus two additional terms. The first of these, $-\left(D^1 - D^0\right) L^1 < 0$, represents the effect of the change in the loss probability. The second one, $[c(x_0) - c(x_1)] > 0$, represents the cost savings due to reduced precaution. If $D_2 = 0$ (i.e., $f_2 = 0$), then there is no externality effect and the combination of these two terms is always positive and simply measures the effect of the individual reoptimizing precaution due to the reduced loss.

However, if $D_2 < 0$, then the first term may dominate the second one (i.e., $-\left(D^1 - D^0\right) L^1 + [c(x_0) - c(x_1)]$ could be negative), so that the true value is less than the naive value. If the externality if is sufficiently large, it can even be the case that $TV < 0$; that is, the LMT decreases welfare because it leads to a large reduction in equilibrium precaution.

Precaution may, of course, also constitute a negative externality (i.e., $D_2 > 0 \Rightarrow f_2 < 0$). For example, as suggested earlier, an individual's probability of falling victim to a home burglary may increase if others invest more in the safety of their homes and thus make them less attractive targets for burglars. In this case, $f_2 < 0$ and $dx/d\theta < 0$ mean that the first term in (13) is positive, implying that the naive valuation will necessarily underestimate the true valuation. Example 1 in the appendix demonstrates explicitly how these various results can come about.

3.2 Probability Reduction Technology

For a probability reduction technology (PRT), an increase in θ reduces the probability of loss, but leaves the size of the loss unaffected (i.e., $L(\theta) = \overline{L}$). The individual's objective function is

$$f(x, \overline{x}, \theta) = -D(x, \overline{x}, \theta)\overline{L} - c(x) \tag{20}$$

The naive valuation of an improvement in θ from θ_0 to $\theta_1 > \theta_0$ is

$$NV \equiv f(x_0, x_0, \theta_1) - f(x_0, x_0, \theta_0) = [D(x_0, x_0, \theta_0) - D(x_0, x_0, \theta_1)]\,\overline{L} \tag{21}$$

which is simply the loss probability reduction, *conditional* on behavior not changing, multiplied by the size of the loss. The true valuation is

$$TV = f(x_1, x_1, \theta_1) - f(x_0, x_0, \theta_0) = NV - [D(x_1, x_1, \theta_1) - D(x_0, x_0, \theta_1)]\,L + [c(x_0) - c(x_1)]\,. \tag{22}$$

Suppose the PRT improvement reduces the productivity of precaution, i.e., $D_{13} > 0 \Rightarrow f_{13} < 0$, then $\frac{dx}{d\theta} < 0$, and further suppose that precaution represents a positive externality ($D_2 < 0 \Rightarrow f_2 > 0$). This is sufficient for an increase in θ to lead to $-[D(x_1, x_1, \theta_1) - D(x_0, x_0, \theta_1)]\, L$ being negative; i.e., $x_1 < x_0 \Rightarrow D(x_1, x_1, \theta_1) > D(x_0, x_0, \theta_1)$.

So the first of the additional terms in the expression for TV (relative to NV) is negative. The second additional term is the reduced cost of precaution, which is positive. Depending on the balance between these two terms, the naive valuation could either over- or under-estimate the true value of the PRT. If the behavioral effect on the loss probability (the first additional term), which includes the magnitude of the externality effect, is strong enough, then the PRT could even lead to a reduction in welfare ($TV < 0$).

Suppose we maintain all of the above assumptions except that now $D_{13} < 0$ ($\Rightarrow f_{13} > 0$). This means that the PRT raises the productivity of individual precaution and so $\frac{dx}{d\theta} > 0$ and $x_1 > x_0$. Consequently, behavioral effects further reduce the loss probability, i.e., $D(x_1, x_1, \theta_1) < D(x_0, x_0, \theta_1)$. The combination of individuals privately reoptimizing their level of precaution and the positive externality of precaution, means that, in this case, the naive valuation necessarily underestimates the true valuation, and the true value must be positive.[10] This, however, is not true if precaution represents a negative externality (i.e., if $D_2 > 0 \Rightarrow f_2 < 0$) as was suggested earlier is a possibility in the case of burglary. Example 2 in the Appendix illustrates all of these possibilities.

4 Discussion and Extensions

We now relate our results to some of the existing research on offsetting behavior and discuss possible extensions to our model. Our goal is to highlight the link between observed behavioral change due to safety technology improvements, and the welfare implications that ensue.

Peltzman (1975) and the welfare effects of offsetting behavior. Much, but certainly not all, of the existing literature either ignores the normative aspects of offsetting behavior or takes an informal approach to welfare analysis. A useful starting point for our discussion is the seminal paper by Peltzman (1975). In the abstract of his paper (p. 677), he writes:

> Technological studies imply that annual highway deaths would be 20 percent greater without legally mandated installation of various safety devices on automobiles. This literature, however, ignores offsetting effects of nonregulatory demand for safety and driver response to the devices. This article indicates that these offsets are virtually complete, so that regulation has not

[10] Although $x_1 > x_0 \Rightarrow c(x_1) > c(x_0)$, the fact that x is chosen optimally means that, on balance, the individual must be better off when θ rises to θ_1.

decreased highway deaths. Time-series (but not cross-section) data imply some saving of auto occupants' lives at the expense of more pedestrian deaths and more nonfatal accidents, a pattern consistent with optimal driver response to regulation.

This passage raises several interesting issues. Firstly, "technological studies" presumably means what we term the naive valuation methodology that ignores changes in behavior. As we emphasize, however, offsetting behavior does not necessarily eliminate the welfare benefits of technological innovation or safety regulation. Even if the accident rate increases and the death rate does not fall as a result of an exogenously imposed safety improvement, it is possible that the regulation has improved welfare, as the cost of precaution is reduced. This is one of the lessons learned from Proposition 3.

Also, example 1 in the Appendix demonstrates that an improved LMT may lead to a welfare improvement that exceeds the naive valuation even if offsetting behavioral change leads to an increase in both the probability of loss and the expected value of loss.[11]

Effects of offsetting behavior on insiders and outsiders. Another aspect of Peltzman's work that suggests a useful extension of our model is that he separates the effects of the safety regulation into changes in the death rate of drivers and pedestrians. Pedestrians do not influence how much precaution, x, is used by drivers. For simplicity, we assume here that drivers don't take the situation of pedestrians into account when choosing precaution.[12] Pedestrians, of course, can be expected to react to changes in the drivers' average level of precaution.

As before, let $f(x, \overline{x}, \theta)$ denote the drivers' utility, and let $x(\theta)$ be their equilibrium level of precaution and $V(\theta) = f(x(\theta), x(\theta), \theta)$ be their value function.

Let $g(y, \overline{x})$ be the utility of a pedestrian when y is the level of care that a pedestrian exerts in order to avoid being involved in an accident with an automobile. Note that there are no interdependencies amongst pedestrians in their choice of y, but they are affected by the average care of drivers. We assume that $g_{11} < 0$ and that an interior optimum for y exists, characterized by $g_1 = 0$. It is natural to assume that $g_2 > 0$; that is, the more careful are drivers the better off are pedestrians.

Because the optimal y depends on x, we can write the equilibrium value of y as $y(\theta) = y(x(\theta))$. This leads to a value function for pedestrians, $\Omega(\theta) = g(y(\theta), x(\theta))$. Using

[11] Of course, as Peltzman (1975, p. 717) notes, it was probably not the intention of the regulators (Congress) that the accident rate should increase as a result of safety regulations regardless of overall welfare effects.

[12] In reality drivers probably do have an incentive to avoid hitting pedestrians, but it seems reasonable that their efforts are directed relatively more to avoid accidents that damage themselves through collisions with other drivers.

the envelope theorem ($g_1 = 0$), $\Omega'(\theta) = g_2 \frac{dx}{d\theta}$. In the case of truly offsetting behavior, $\left(\frac{dx}{d\theta} < 0\right)$, which applies for any improved LMT of the type described in section 3.1, it follows that $\Omega' < 0$. Given a population made up of a fraction λ of drivers and $(1-\lambda)$ of pedestrians, the overall value function is $\Psi(\theta) = \lambda f(x(\theta), x(\theta), \theta) + (1-\lambda) g(y(\theta), x(\theta))$. The welfare effect of a mandatory marginal increase in car safety is

$$\Psi'(\theta) = \lambda \left[f_3 + f_2 \frac{dx}{d\theta} \right] + (1-\lambda) g_2 \frac{dx}{d\theta} \tag{23}$$

Given truly offsetting behavior, the inclusion of pedestrian interests make it less likely that an improvement in safety technology increases welfare.

Consider, however, the case of an improved safety technology such as anti-lock brakes that, ceteris paribus, reduce the probability of an accident (i.e., a PRT). If the innovation also enhances the marginal benefit of precaution, then $f_{13} > 0 \Rightarrow \frac{dx}{d\theta} > 0$, and precaution exhibits a positive external effect ($f_2 > 0$), it follows that the first term of $\Psi'(\theta)$ is positive and $g_2 > 0$ implies that the second term is also positive. Under these conditions, including pedestrians' utility in welfare considerations actually enhances the value of the improved safety technology.

Many recent advances in car safety technology appear to possess both loss mitigation and probability reducing effects. It is not obvious whether these raise or lower the marginal benefit of precaution. A proper welfare analysis requires that this be sorted out.

Our example of pedestrians as a group not directly involved with the loss prevention activities, but affected by them, can be generalized to other environments. Consider a first group of "insiders" (the drivers in our example) and a second one of "outsiders" (the pedestrians in our example).[13] In many scenarios, including the interest of outsiders is important for welfare analysis.

Consider, for example, a pair of individuals who have an ongoing sexual relationship. Suppose one engages in high risk sex with multiple partners while the other treats the relationship as monogamous. The former can be treated as an insider who chooses precaution x to reduce the chances of incurring a sexually transmitted disease (STD).

Suppose the result is an equilibrium in which an improved treatment for any STD generates higher risk behavior by insiders that leads to a higher rate of infection in this group. Although a LMT such as improved treatment in case of infection may improve the expected wellbeing of insiders, outsiders may well be subjected to a worsening in risk and welfare, and this reduces the chances of a LMT improving overall welfare. Note that this conclusion follows whether or not outsiders can make (costly) choices that can counteract the higher risk they face. The result follows as long as $g_2 > 0$, whether or not $g_1 = 0$.

[13] Outsiders may be created by institutions. For example, in countries with public health care plans, outsiders in the context of traffic safety could be taxpayers who incur the cost of healthcare for those in accidents.

Other related literature. As noted in our introduction, much of the literature about offsetting behavior is directed at determining empirically its size in a wide variety of economic settings. Development of general theoretical models for the purpose of background and welfare assessment has lagged. Over two decades ago, Neill (1993, p. 435) pointed out that "The controversy over the effectiveness of safety regulation in decreasing accident rates is now decades old. However, the arguments that have been made are, for the most part, empirical. The lack of theoretical perspective [...] makes it impossible to predict when and what sort of regulation is likely to be effective."

Our model provides such a theoretical framework for welfare analysis in settings where offsetting behavior is important and likely influenced by the state of technology. Some of the elements of our model are present in existing theoretical models of offsetting behavior, 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 analyze, 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. The externality arises when airlines accept the baggage of travelers who are transferring from other airlines without rechecking. An airline's ability to avoid bad outcomes is affected if other airlines do not adopt a secure baggage checking system. 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 focuses on 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.

Muermann and Kunreuther (2008) incorporate some of the elements of Kunreuther and Heal (2003) to model optimal investment in self-protection in the presence of interdependencies with others' decisions (i.e.; "contamination"). In their model, an individual's effort affects his probability of an "own fault" accident. 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 externality in their model, as each agent's effort level 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 welfare improvement.

Gossner and Picard (2005) investigate how to value the benefit of a road safety improvement when offsetting effects are present. In their model, the loss is financial, and externalities work through the insurance market rather than through some explicit interaction effects as in our model or in the papers discussed above. They do, however, consider a similar problem as in our paper by taking into account how changes in road safety affect individuals' precautionary effort levels. Because losses are financial, they also investigate how drivers' risk aversion affects the value of road safety improvements.

Hause (2006) is restricted to the case in which an individual's precaution affects the loss probability while the exogenously introduced safety technology affects the loss size (i.e., our LMT model). His model does not include an externality effect – be it positive or negative – or interaction effect of others' safety choice and so, while his approach to determining the difference between the naive value and the true value of an exogenous improvement to safety is similar to ours, his welfare conclusions only apply for an environment in which these factors are irrelevant.

Besides these models developed for specific contexts, there are somewhat more general models of offsetting behavior. Our model should be thought of as further developing those models. Of particular relevance is Blomquist (1986) who develops a general model of driver safety behavior in which an individual's own effort and a measure of exogenous safety jointly affect both the probability of an accident and the size of loss. Therefore, his model incorporates aspects of both LMT and PRT. However, he does not model any externality effect and so limits the use of his model for welfare analysis since, as we have demonstrated, the presence, sign, and size of the externality effect affect the social value of an improved safety technology. His analysis also ignores the possibility that an exogenous safety improvement may increase the marginal product of an individual's own effort; i.e., he effectively assumes $f_{13} < 0$ and so only deals with the case that $\frac{dx}{d\theta} < 0$.

Neill (1993) develops a model of exogenous safety improvements through regulations that, ceteris paribus, either reduce the probability of an accident (PRT) or reduce the size of loss (LMT). Individuals choose a level of precaution, as well as a private expenditure that reduces the size of the loss. His goal is to determine conditions on the technological relationships of these variables that allow one to sign the effect of improved exogenous safety on the equilibrium probability of an accident. His results are similar to ours in that whether a safety improvement increases or decreases the marginal benefit of precaution is crucial for whether the net effect is a reduction in the equilibrium probability of accident.[14] As with the model of Blomquist (1986), his model does not include an externality effect, which leads to the same limitations for welfare analysis as explained above.

[14]However, he has the added complication in regards to how expenditure on self-insurance is affected.

5 Conclusions

We have presented a general model of offsetting behavior that provides positive and normative analysis of either exogenous (mandated) or endogenous (voluntarily chosen) safety improvements. Our model combines an explicit treatment of (1) strategic behavior, including consideration of whether precaution levels are strategic complements or substitutes, (2) an externality effect between people's activity levels, be it positive or negative, (3) how the safety technology affects the marginal value of precaution, (4) whether the level of safety technology is exogenously imposed or endogenously chosen. We also develop specialized versions of our model in which the technological improvement reduces either the size or the probability of the loss, but not both.

Most earlier models analyzed the positive implications of offsetting behavior, in particular, the extent to which OB offsets the direct reduction in exposure to the loss due to the imposed safety technology. In contrast, a central focus of our analysis is normative, and we derive conditions for the net welfare effect of the technology improvement to be positive or negative.

As we show, on its own, the size and direction of the offsetting behavior effect is neither a useful indicator for whether an improved safety technology improves or worsens welfare, nor for the size of this welfare effect. We also investigate offsetting behavior when different technology levels, such as the range or quality of optional safety features for automobiles, are available at private cost to individuals. Whether it is welfare enhancing to tax or subsidize the technology depends on whether the technological improvement enhances or reduces the marginal benefit of each individual's own level of precaution in conjunction with whether the relevant externality is a positive or negative one. Moreover, the size of the optimal tax depends on measures 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 without giving agents the possibility to insure against losses, it would certainly be interesting to analyze some of the questions and policies of those papers within our model framework.

One can also 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 behavior.

Acknowledgments

We thank seminar participants at the University of Guelph, University of Georgia, ETH Zurich, University of Montreal, and CORE for useful comments. In particular, we thank James Amegashie, Antoine Bommier, Christian Gollier, Richard Peter, Pierre Picard, Casey Rothschild, and Art Snow for very useful suggestions. The first author's research is funded in part by SSHRC.

References

Ayres, I. and S. D. Levitt (1998): "Measuring Positive Externalities From Unobservable Victim Precaution: An Empirical Analysis of Lojack", *The Quarterly Journal of Economics*, vol. 113, no. 1, pp. 43-77.

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.

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.

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.

Hammitt, James K. (2005): "Methodological Review of WTP and QALY Frameworks for Valuing Environmental Health Risks to Children", *OECD, Project VERHI*, pp. 29.

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.

Hirth, R. A., M. E. Chernew, E. Miller, A. Mark Fendrick, William G. Weissert (2000): "Willingness to Pay for a Quality-adjusted Life Year: In Search of a Standard", *Medical Decision Making*, vol. 20, no. 2, pp. 332-342.

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 Safety 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 behavior and the 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.
S. Shavell (1991): “Individual Precautions to Prevent Theft: Private versus Socially Optimal Behavior”, *International Review of Law and Economics*, vol. 11, pp. 123-132.
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.

6 Appendix

6.1 Example 1: LMT

Consider the following specific example of the LMT model above. Let the loss probability be

$$D(x, \overline{x}) = \kappa - bx - (1-b)\overline{x} \tag{24}$$

so that b (with $0 \leq b \leq 1$) measures the extent to which the loss probability is influenced by the individual's own level of care, while $(1-b)$ measures the extent to which the loss probability is influenced by the average level of care of other players. Let the cost of precaution be $c(x) = \alpha x^2$ and let the loss function be $L(\theta) = \overline{L} - \theta$. Note that smaller b implies own effort is less important relative to others effort. The objective function, equation (15), becomes:

$$f(x, \overline{x}, \theta) = -\left[\kappa - bx - (1-b)\overline{x}\right]\left(\overline{L} - \theta\right) - \alpha x^2 \tag{25}$$

Individual optimization yields $x = \frac{bL}{2\alpha}$ which also, of course, is the level of precaution of everybody else in equilibrium.[15] Substituting this value for x (and $\overline{x}$) into f and simplifying yields the value function:

$$V(\theta) = -\kappa L(\theta) + \frac{b(2-b)\left[L(\theta)\right]^2}{4\alpha} \tag{26}$$

The following numerical examples illustrate the possible differences between the naive and true valuations of discrete safety improvements. Letting $\kappa = 1.2$, $\alpha = 5$, $\overline{L} = 10$, we evaluate both NV and TV for a safety improvement from $\theta_0 = 0$ to $\theta_1 = 1$. Note that $x = \frac{bL(\theta_0)}{2\alpha}$ and $f_2 = (1-b)L(\theta) = (1-b)(10-\theta) \geq 0$ since $0 \leq b \leq 1$.[16] Choosing $b = 0.85$ generates values $NV = 0.35$ and $TV = 0.27$; that is, the naive valuation exceeds the true valuation. Alternatively, choosing $b = 0.98$ generates the values $NV = 0.22$ and $TV = 0.25$ which demonstrates that, even in the presence of offsetting behavior with an externality effect, the true value of a safety improvement may exceed the naive value.

Choosing $\kappa = 0.4$ and maintaining all of the other parameter values, we find that for b small enough, it is possible to generate cases where the true value of a safety improvement is negative; that is, the offsetting effect coupled with a strong enough externality effect leads to a worsening of welfare. In particular, $b = 0.3$ generates values $NV = 0.1$ but $TV = -0.084$.

[15] For an interior solution to apply, we have to restrict parameters such that the probability of loss lies between zero and one: $\frac{bL}{2\alpha} < \kappa < 1 + \frac{bL}{2\alpha}$.

[16] This implies that precaution of others represents a positive externality except for $b = 1$ which implies there is no externality whatsoever.

Finally, note that we can generate an example with a negative externality (i.e., $D_2 > 0 \Longrightarrow f_2 < 0$) by simply replacing the term $-(1-b)\overline{x}$ with $+a\overline{x}, a > 0$ in equation (24). In this case TV is assured to exceed NT.

6.2 Example 2: PRT

In our first version of a PRT, we let $D(x, \overline{x}, \theta) = [\kappa - b\theta x - (1-b)\theta\overline{x}]$, $b \in [0,1]$. In this case, an increase in θ not only reduces the probability of loss but also increases the effectiveness of individual precaution; that is, for any given level of x, an increase in θ increases the marginal benefit of being careful. Specifically, $D_{13} = -b$ and so $f_{13} = b\overline{L} > 0$. Each individual minimizes

$$f(x, \overline{x}, \theta) = -[\kappa - b\theta x - (1-b)\theta\overline{x}]\,\overline{L} - \alpha x^2. \tag{27}$$

We have $f_2 = (1-b)\theta\overline{L} > 0$ and so the externality is positive for $b \in [0,1)$. Solving the first order condition yields $x = \frac{b\theta\overline{L}}{2\alpha}$, which is also equal to the equilibrium level of care by everybody else. Note that $f_{13} > 0$ for $b \in (0,1]$ implies $\frac{dx}{d\theta} > 0$; i.e., an increase in θ increases the equilibrium level of care which is the opposite of the usual idea of (truly) offsetting behavior.

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

$$V(\theta) = -\left[\kappa - \theta\frac{b\theta L}{2\alpha}\right]\overline{L} - \alpha\left(\frac{b\theta\overline{L}}{2\alpha}\right)^2. \tag{28}$$

Differentiating V with respect to θ and noting that $0 < b < 1$ yields

$$V'(\theta) = -\frac{b\overline{L}^2}{\alpha}\left(\frac{b}{2} - 1\right)\theta > 0. \tag{29}$$

Suppose, however, that we make a similar alteration as we did for the base LMT model above (i.e. replace the term $-(1-b)\theta\overline{x}$ with $+a\theta\overline{x}, a > 0$). This implies a negative externality in precaution which, in conjunction with $\frac{dx}{d\theta} > 0$, creates a negative welfare effect. If this effect is sufficiently strong, it could lead to an improved PRT actually worsening welfare.

In contrast, the second example of a PRT illustrates a case in which an increase in θ that reduces the loss probability leads to people being less careful in equilibrium. Assume now that $D(x, \overline{x}) = (1-\theta)[\kappa - bx - (1-b)\overline{x}]$, where $0 < \theta < 1$. An interpretation of this function is an immunization or a treatment that is effective only for a proportion θ 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 maximizes

$$f(x, \overline{x}, \theta) = -(1-\theta)\left[\kappa - bx - (1-b)\overline{x}\right]\overline{L} - \alpha x^2. \tag{30}$$

This yields $x = \frac{b(1-\theta)\overline{L}}{2\alpha}$, which also equals the equilibrium level $\overline{x}$. For the initial value $\theta = \theta_0$, we have $x_0 = \frac{b(1-\theta_0)\overline{L}}{2\alpha}$. For the range $\theta \in (\theta_0, \theta_1)$, we have $f_{13} < 0 \Longrightarrow \frac{dx}{d\theta} < 0$, while $f_2 = (1-\theta)(1-b)\overline{L} > 0$. Since $\frac{dx}{d\theta} < 0$, the positive externality of precaution in conjunction with the offsetting effect leads to a negative effect on welfare due to the improved safety technology. If this effect is strong enough, the naive valuation will overestimate the true valuation which could even be negative.

We can see these possibilities by solving for the value function. For this example to have an interior solution, we have to assume that $\kappa \geq \frac{b(1-\theta)\overline{L}}{2\alpha}$ (to ensure that the equilibrium loss probability is non-negative). Substituting the equilibrium values of x and $\overline{x}$ in the objective function and simplifying yields

$$V(\theta) = -\kappa(1-\theta)\overline{L} + \frac{b(1-\theta)^2\overline{L}^2}{4\alpha}(2-b). \tag{31}$$

We can differentiate this with respect to θ to yield

$$V'(\theta) = -\overline{L}\left[\frac{b(1-\theta)\overline{L}}{2\alpha}(2-b) - \kappa\right]. \tag{32}$$

Using the restriction on κ from above, one can see that this expression may be negative if b is smaller than 1, indicating that welfare may indeed fall with an increase in θ. If, instead, $b = 1$, then the restriction on κ guarantees that $V'(\theta) > 0$, which implies that an increase in θ is beneficial for individuals. This must be the case since $b = 1$ means there is no externality involving precaution of others.