Peeking into the Black Box: Technological Transparency and its Impact on Self-Protection

February 15, 2019

Abstract

Technological transparency refers to perfectly understanding the joint determination of a risky outcome by nature and the decision-maker's action. This paper discusses the impact of technological transparency on self-protection (loss prevention). When combined with ex ante observable exogenous risk factors, technological transparency creates value by always inducing the most efficient preventive effort. When the exogenous risk factors are only ex post observable, technological transparency allows the decision-maker to objectively attribute the (non-)occurrence of the loss to his own effort and external causes in hindsight, which provides a channel for regret aversion to raise self-protection. The results highlight the positive value of understanding causal determinants of risks in promoting adequate and efficient preventive effort.

Keywords: technological transparency · self-protection · regret aversion · value of information

JEL-Classification: D61 · D81 · D83 · D91 · O33

1 Introduction

Most life outcomes depend inevitably on both our own actions and various factors beyond our control. In the context of self-protection (also referred to as loss prevention, see Ehrlich and Becker, 1972), which is a costly effort to reduce the likelihood of a loss event, the success of preventive effort usually depends on exogenous determinants of the underlying risk. The HPV vaccine, for instance, targets only a few out of nearly two hundred types of viruses and is most effective in preventing cervical cancer only for women within a certain age range. In disaster prevention, a small effort to reinforce a house may help the house withstand a moderate hurricane, whereas a large effort may fail to do so if the hurricane turns out to be catastrophic. While healthy diet and regular physical exercise help reduce the risk of diabetes, the actual efficacy of lifestyle change in preventing diabetes is shown to depend on one's genotype. The mechanism of the interaction between effort and genes is, however, not yet perfectly understood (Marx, 2002).

Same as all aforementioned examples, any self-protection technology has an inherent possibility of failing. While a decision-maker always associates more effort with a lower loss probability (equivalently, a lower probability of failing), he still lacks knowledge about the exogenous determinants of the risk that jointly predict the actual success of his effort. In self-protection models, this limited knowledge is holistically reflected by a loss probability lying strictly between 0 and 1. Meanwhile, the causes of the effort's potential failure have received little discussion. In other words, the underlying mechanism of self-protection resembles a black box in its conventional interpretation.

Technological transparency (TT) refers to fully understanding the joint determination of a risky outcome by nature and the decision-maker's action. When applied to self-protection, TT corresponds to opening the black box by revealing the exogenous risk factors together with how they collectively predict the success of the preventive effort. Mapping the actual intensity of a hurricane to the minimum amount of reinforcement to keep a house safe is an example of TT when the intensity of the hurricane is the only exogenous risk factor. In disease prevention, TT may take the form of understanding the dependency structure between the efficacy of preventive healthcare and an individual's genetic information, which, for diabetes prevention, is shown to be associated with more than 100 DNA variations (see Marx, 2002).

This paper investigates the impact of TT on the demand for self-protection. For this purpose, it is crucial to distinguish between two types of exogenous risk factors that are observable either before or after the resolution of uncertainty, i.e. when the decision-maker witnesses the actual (non)occurrence of the loss.

When combined with *ex ante* observable risk factors, TT corresponds to possessing perfect information à la Gollier (2001) and therefore raises the decision-maker's welfare by inducing the most efficient effort conditional on any realized combination of the risk factors. In the paper, we flesh out the positive value of TT by demonstrating three fractions of losses for a given self-protection technology: those that are only prevented under TT, those that are always prevented but only at the lowest cost under TT, as well as unpreventable losses that would have effort completely wasted without TT. We also derive sufficient conditions on the probability function for the overall value of TT to increase or decrease. As a direct consequence of TT's positive value, before making his decision, a rational decision-maker is always incentivized to reveal the values of the ex ante observable exogenous risk factors, for instance by conducting a genetic test in case his genotype determines the success of his preventive effort[1]. This result is in contrast with its counterpart in insurance markets with symmetric information about the insured risk, where the policyholder chooses rather not to know their genotype for fear of the premium risk ((see Crocker and Snow, 2000; Peter et al., 2016)).

In other situations, the exogenous risk factors may be only *ex post* observable. For instance, one can only obtain the actual intensity of a hurricane after its occurrence. In such cases, TT allows the decision-maker to realize what he should have done in hindsight. While the hindsight does not affect an expected utility maximizer's ex-ante decision, it can be crucial when the decision-maker's utility depends on a reference point taking the form of a counterfactual outcome. A well-known case is when the decision-maker exhibits regret aversion and anticipates future disutility from realizing having made a suboptimal choice (see for instance in Bell, 1982; Loomes and Sugden, 1982; Quiggin, 1994). Together with ex post observable risk factors, TT allows the decision-maker to objectively attribute the (non-)occurrence of a loss to himself and to external causes in retrospect. We show that this objective attribution allows regret aversion to induce ex ante more self-protection. Prior literature may suggest regret aversion leads to "less extreme" decisions by either raising or lowering the optimal decision variable, as shown by Braun and Muermann (2004) for insurance demand and by

[1] Assume for simplicity that such tests are available at no cost.

Muermann et al. (2006) for portfolio choice. Our result shows that the same conclusion may not easily generalize to other types of decisions. In particular, our results highlight the fundamental difference between two broad types of risk management activities: (self-)insurance and self-protection (see Courbage et al., 2013, for a recent survey). Since the seminal work by Ehrlich and Becker (1972), it is well-known that while insurance and self-insurance are substitutes, insurance and self-protection may be either substitutes or complements. Furthermore, risk aversion is known to raise the demand for insurance and self-insurance, but not necessarily the demand for self-protection (see for instance Dionne and Eeckhoudt, 1985; Briys and Schlesinger, 1990). In this paper, this difference is again reflected by the impact of regret aversion on the demand for both types of activities.

Our findings suggest that more efficient and adequate risk reduction may be enabled by scientific research uncovering exogenous risk factors and their interaction with preventive effort. The rest of this paper is structured as follows. Section 2 defines TT and interprets it in selected application contexts. Section 3 demonstrates TT's positive value when combined with ex-ante observable risk factors. Section 4 shows that under TT and ex-post observable risk factors, self-protection increases with regret aversion. Section 5 addresses cases where TT is unavailable or not fully available. Section 6 concludes and discusses the implications of our results.

2 Technological Transparency: inside the Black Box of Self-Protection

Consider a decision-maker with the initial endowment $w > 0$ who faces a potential loss L. L is a constant whose value lies strictly between 0 and w. Self-protection means the decision-maker may engage in some costly effort in exchange for a reduction of the loss probability. The loss probability $p(x)$ is assumed to be twice differentiable, decreasing and convex: $p' < 0, p'' > 0$, where x is the amount of disutility induced by the effort that is additively separable from the utility of consumption [2]. Under these assumptions, the choice of the optimal self-protection

[2] The cost of preventive effort may be perceived as separable from the utility of consumption for many reasons. In disease prevention, the effort often takes the form of healthy diet, regular physical exercise, lower cigarette and/or alcohol consumption, etc. Engaging in such activities are usually not associated with lower consumption. Another typical case is when the effort significantly precedes the resolution of uncertainty, which leads to a two-period self-protection problem (see Menegatti, 2009, for instance), where the cost – either monetary or not – is also separated from the consumption in the second period.

effort corresponds to the following optimization problem:

$$\max_{x \geq 0} U(x) = [1 - p(x)]u(w) + p(x)u(w - L) - x, \tag{1}$$

whose solution, denoted by x^*, is determined by the first order condition:

$$-p'(x^*)[u(w) - u(w - L)] = 1. \tag{2}$$

The left hand side of Equation 2 represents the expected marginal benefit and the right hand side represents the marginal cost of an additional unit of effort. Equation 2 can also be written as:

$$-p'(x^*) = \frac{1}{\hat{x}}, \tag{3}$$

where $\hat{x} = u(w) - u(w - L)$ corresponds to the highest effort the decision-maker is willing to undertake knowing the loss will happen for sure. Note that $\hat{x}$ also represents the utility premium induced by the loss (Friedman and Savage, 1948). The second order condition of 1 is automatically fulfilled given the aforementioned assumptions. As depicted in Figure 1, x^* is the point where the tangent line along the $p(x)$ curve is parallel to the line connecting $(0, 1)$ and $(\hat{x}, 0)$.

An inherent property of self-protection is that the effort – no matter how large – may either succeed or fail [3]. While prior self-protection models assume $p(x)$ to be decreasing and convex with almost no exception, they are silent regarding when and why the effort will succeed or fail. We aim to discuss exactly the determinants of the success of self-protection while preserving the properties of $p(x)$ commonly adopted by earlier studies. To do so, let $\tilde{Y} = (\tilde{y_1}, \tilde{y_2}, \tilde{y_3}, ..., \tilde{y_n})$ denote the vector of exogenous risk factors. $\tilde{Y}$ contains $n \geq 1$ random variables whose values are unknown at the time the decision is made. Assume $\tilde{Y}$ is an exhaustive collection of relevant risk factors so that each realization of $\tilde{Y}$ corresponds to one particular state of the world where any effort either succeeds or fails with absolute certainty.

Consider, for instance, that the decision-maker wants to reinforce his house to prevent it from being destroyed by a hurricane. For simplicity, assume the success of his effort depends solely on the intensity of the hurricane $\tilde{y}$ so that $n = 1$ and $\tilde{Y} = \{\tilde{y}\}$. Each given intensity

[3] I assume $p(x)$ to be always strictly between 0 and 1. However, all conclusions in this paper remain unaffected if $p(x) = 1$ or $p(x) = 0$ are allowed for some effort levels as long as $p' < 0$ continues to hold.

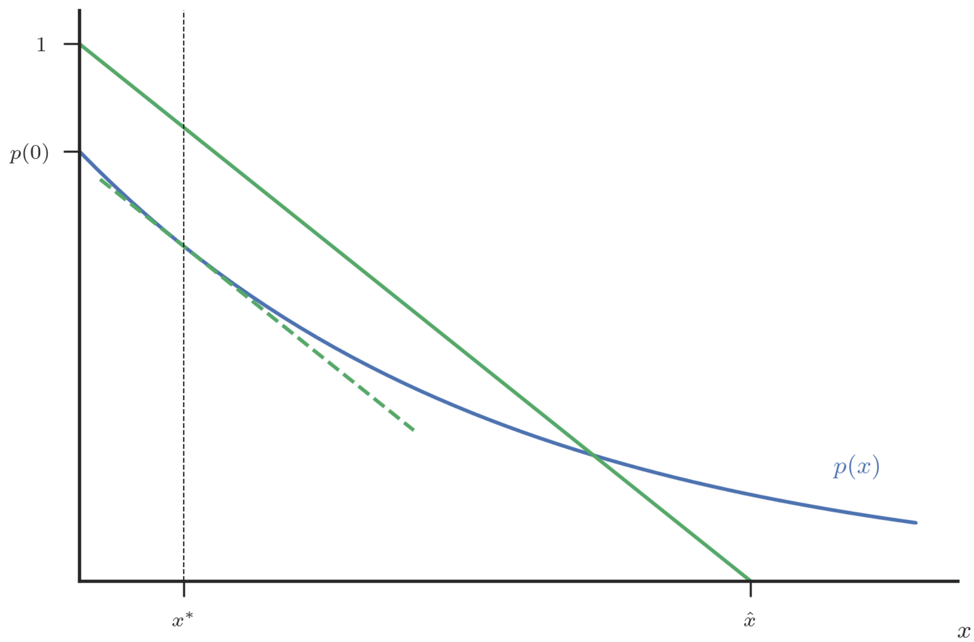

Figure 1: Optimal self-protection without technological transparency. x^* is the optimal effort. $\hat{x} = u(w) - u(w - L)$ is the utility premium induced by the loss.

y, or each realization of $\tilde{y}$, requires a minimum effort $t = \theta(y)$ such that the house is kept safe. The higher the intensity, the larger the minimum required effort. Assume further that when choosing his effort, the decision-maker knows the distribution of $\tilde{y}$ as well as the function θ so that he perceives the minimum effort to succeed as a random variable $\tilde{t} = \theta(\tilde{Y})$ and the loss probability as the probability of this random variable exceeding his effort: $p(x) = Prob\{x < \tilde{t}\}$.

In the examples above, the reduction of the loss probability by self-protection is reinterpreted in terms of the (joint) distribution of the exogenous risk factors. Note that this reinterpretation is only possible because the decision-maker knows the function θ that maps the realization of the exogenous risk factors to the minimum effort to successfully prevent the loss. Also note that this reinterpretation makes the underlying mechanism of self-protection transparent without changing $p(x)$ per se. Let us now formally describe the ideas above with the following two definitions.

Definition 1. In self-protection, the *conditional threshold effort* is the lowest effort such that the loss does not occur. Mathematically, it is represented by the random variable $\tilde{t}$ satisfying $p(x|\tilde{t}) = \mathbb{I}\{x < \tilde{t}\}$, $\mathbb{I}$ being the indicator function.

In other words, knowing the conditional threshold effort means perfectly predicting the success/failure of any effort. The concept of the conditional threshold effort rests on the straightforward intuition that in each state of the world, as long as some effort successfully prevents the loss, any effort exceeding it will also successfully prevent the loss. When interpreting self-protection merely as a reduction of the loss probability, the states are latent and the state space is partitioned into two events, or "possibility sets" (see Aumann, 1976) [4]: loss and no loss. Since the states within either event are not distinguishable from one another, the mechanism of self-protection is hidden in its conventional interpretation. Definition 1 forces the states to emerge by allowing them to be fully represented by the conditional threshold effort. A larger effort reduces the probability of the loss event by exceeding more potential realizations of the conditional threshold effort and therefore pushing more states into the no loss event.

The following proposition addresses the distribution of the conditional threshold effort.

Proposition 1. *Given a self-protection technology characterized by the probability function $p(x)$ with $p' < 0, p'' > 0, 0 < p < 1$. The conditional threshold effort $\tilde{t}$ follows a mixed type distribution with the cumulative distribution function $F(t)$ such that:*

- $F(t) = 0,\qquad \text{if } t < 0$
- $F(t) = 1 - p(t) = 1 - p(0) + \int_0^t f(u)\mathrm{d}u,\qquad \text{if } t \geq 0,$

where $f(t)$ corresponds to the probability density function of $\tilde{t}$. Furthermore, it holds that:

- $f(t) = -p'(t),\qquad \text{if } t > 0$
- $f(t) = 0,\qquad \text{if } t < 0$
- $f(0)$ *is non-existent.*

We leave all mathematical proofs into the appendix. The insight offered by Proposition 1 is that the entire distribution of the conditional threshold effort can be derived from $p(x)$. It is worth emphasizing though that for any self-protection technology, it is its mechanism $F(x)$ that *determines* its effectiveness $p(x)$ in the first place. In reality, however, we usually first get to know $p(x)$ from experience, which, if based on long enough and large enough observations,

[4] see an application of the same concept by Neumuller and Rothschild (2017) on limited financial literacy.

should yield an unbiased estimation. Proposition 1 therefore infers $F(t)$ backwards from $p(x)$. Since the purpose of this paper is not on the biasedness of $p(x)$ but on the underlying mechanism of an arbitrary self-protection technology represented by any $p(x)$, we shall abstain from discussing the former and always treat $p(x)$ as a predetermined, unbiased function that is common knowledge a priori.

So far, Definition 1 reveals the existence of the states, whose distribution is derived by Proposition 1 for a given self-protection technology. It remains open when and how a decision-maker can find out the true state of the world.

Definition 2. In self-protection, *technological transparency* (TT) means the ability to identify the conditional threshold effort through the observation of exogenous risk factors. Mathematically, it means the existence of a function θ such that $\tilde{t} = \theta(\tilde{Y})$.

TT helps the decision-maker identify the true state upon the observation of the exogenous risk factors. When translated to applications, TT means fully understanding the joint determination of the risky outcome by nature and the decision-maker's preventive effort. Its achievement generally involves two steps: identifying the relevant exogenous risk factors and estimating the conditional threshold effort as a function of those risk factors. The hurricane example mentioned earlier in this section is a simplified setting where the intensity of the hurricane is the only exogenous risk factor. Hence, TT is equivalent to knowing the function θ that maps each potential intensity of the hurricane to its conditional threshold effort. In more complex applications such as disease prevention, θ may have multiple input variables. TT is usually achieved through conducting well-designed controlled experiments, observational studies, or simulations. In some cases, TT may also be approximated through experiential or social learning. In the following two sections, we shall explore the impact of TT on the demand for self-protection depending on when the exogenous risk factors are observed. Section 5 offers a brief discussion on cases where TT is unavailable or not fully available.

3 Ex Ante Observables and the Value of Information

In this section, we investigate the consequence of TT in combination with exogenous risk factors that are ex ante observable. We start with the benchmark scenario described at the beginning of Section 2 where the decision-maker only knows $p(x)$. Now assume both TT and the values of the exogenous risk factors are available. How will this affect the decision-maker's optimal effort?

Proposition 2. *Under TT, when the exogenous risk factors are observed ex ante, the optimal self-protection effort equals the conditional threshold effort if the conditional threshold effort does not exceed the utility premium induced by the loss; it equals zero otherwise.*

The intuition underlying Proposition 2 is straightforward: Together with known values of the exogenous risk factors, TT reveals the true conditional threshold effort before the decision-maker makes his choice. Hence, whenever the loss is preventable, the decision-maker chooses to prevent the loss at the lowest possible cost [5]. If the loss is unpreventable instead, he chooses not to exert any effort. Note that we use the word unpreventable to refer to two situations: either the conditional threshold effort is infinite or it is finite but "not worthwhile", i.e. it costs more than the gain from avoiding the loss. Since both situations have the same behavioral consequence, we do not distinguish between them for the sake of convenience.

Depending on how the conditional threshold effort compares to the optimal effort x^* in the benchmark case, TT may make the decision-maker better off in three different ways, which we summarize in the following corollary.

Corollary 1. *Consider a self-protection technology characterized by the probability function* $p(x)$ *with* $p' < 0, p'' > 0, 0 < p < 1$. *TT in combination with ex ante observable exogenous risk factors makes the decision-maker better off in one of the three following ways:*

(a) With probability $1 - p(x^*)$, *a preventable loss is prevented both with and without TT. However, its prevention costs less with TT than without TT.*

(b) With probability $p(x^*) - p(\hat{x})$, *a preventable loss is prevented with TT but would occur without TT.*

(c) With probability $p(\hat{x})$, *the loss is unpreventable. No effort is spent with TT whereas effort is wasted without TT.*

In fact, TT and the exogenous risk factors collectively correspond to a conclusive information structure. Such an information structure is known to fully remove risk from the decision

[5] We restrict ourselves to a one-shot decision problem where the loss either does or does not occur exactly once.

problem by providing perfect information, which is valuable for the decision-maker by always inducing the best possible choice (see Gollier, 2001; Blackwell, 1951). Corollary 1 describes the specific channels (together with their corresponding probabilities) for the information to generate non-negative value in the self-protection problem. In particular, when we interpret $p(x)$ as the average loss probability based on the entire population, each of the probabilities in Corollary 1 (a) - (c) represents a specific fraction of the population that benefit from TT in their own way. While the number of deaths due to insufficient prevention is often used to emphasize the importance of preventive healthcare (see for instance Danaei et al., 2009), Corollary 1 (b) suggests that this number may be partly reduced by TT as long as exogenous risk factors are immediately observable (such as gender, age, medical history, etc.). However, in addition to leaving no more preventable losses unprevented, TT may also generate value through saving costs that would be otherwise wasted independent of the loss being preventable or not. The next proposition summarizes the overall value of TT considering all three channels together.

Proposition 3. *When combined with ex ante observable risk factors, the value of TT is always positive for a self-protection problem. Furthermore, this value increases when the technology undergoes a global improvement that preserves x^* and $p(x^*)$.*

The first part of Proposition 3 is a natural consequence of the non-negative value of information applied to self-protection [6]. However, the comparative statics for the value of information is generally not straightforward. As shown by Gould (1974) and Hilton (1981) among others, unlike intuition might suggest, the value of information does not increase with risk aversion, the riskiness of the prior distribution, or the amount of transmitted information (measured by entropy, see Shannon, 1948). Proposition 3 identifies a condition related to technological improvement (see Hoy and Polborn, 2015; Lee, 2015; Li and Peter, 2018) that leads to a higher value of TT: Whenever there is a global improvement of the self-protection technology in that the loss probability becomes lower for any effort other than x^*, the value of TT becomes larger. Note that this condition is equivalent to a deterioration of the conditional threshold effort in the sense of first order stochastic dominance, which can be easily shown with the help of Proposition 1. We illustrate Proposition 3 in Figure 2, where the value of

[6] Note that in the proof of Proposition 3, we follow Gollier (2001) and use the value of information defined in utility terms, but the conclusion is unaffected by the measurement of the value of information (see alternative measures in Hilton, 1981).

TT is represented by the joint area of the three shaded regions (see the proof of Proposition 3 in the appendix). Interestingly, without TT, the technological improvement described in Proposition 3 has neither behavioral nor welfare consequence at all. This is because the optimal effort without TT is determined solely by the local condition described by Equation 3, which completely ignores the improved outcome for every effort other than x^*. However, under TT, the global improvement is fully reflected by an increased welfare through an FSD reduction of the conditional threshold effort.

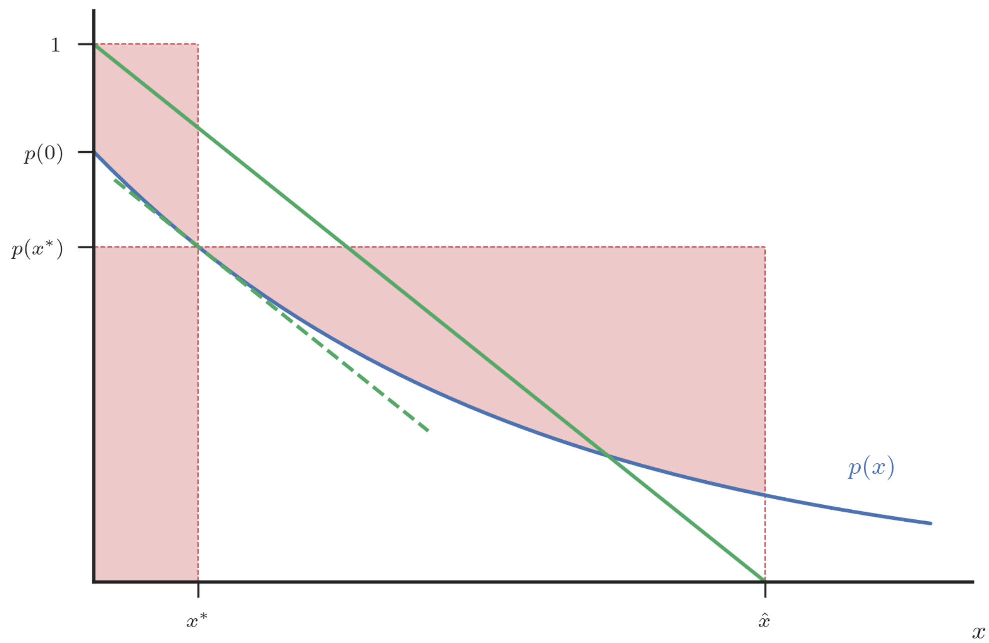

Figure 2: The value of TT in combination of ex ante observable exogenous risk factors. x^* is the optimal effort without TT. $\hat{x} = u(w) - u(w - L)$ is the utility premium induced by the loss. The total area of the shaded regions equals the value of TT.

4 Ex Post Observables and the Impact of Regret Aversion

In many applications, the exogenous risk factors may be observable only after the choice is made. For instance, one only knows the exact intensity of the next hurricane or earthquake after the occurrence of the event. In a penalty kick of a soccer game, the goalkeeper only realizes his opponent's strategy at the end of the kick (And the same works for the kicking player as well.). In these situations, it is no longer possible to remove the risk ex ante. We show in this section that even when the decision-maker faces the exact same risk, TT still has its impact on the optimal effort through anticipated future regret.

There is ample evidence showing people are not always forward looking as assumed by classic decision theories such as the expected utility theory. Regret is a concept well documented by psychologists since more than a century ago (see Zeelenberg and Pieters, 2007, for a survey). Regret theory, which is initially discussed in the economic literature by Bell (1982); Loomes and Sugden (1982); Fishburn (1982), assumes people experience disutility from realizing having made a suboptimal choice (see Bleichrodt et al., 2010; Camille et al., 2004, for existing empirical support for regret theory). While the initial regret theories are restricted to decision problems with 2 alternative actions, Sugden (1993) and Quiggin (1994) axiomize the theory and generalize it to arbitrary choice sets. More recently, regret theory has been applied to various economic decisions including insurance demand (Braun and Muermann, 2004) and portfolio choice (Muermann et al., 2006; Muermann and Volkman Wise, 2006) and is shown to explain observed deviations from predictions of the expected utility theory including the Allais paradox, preference for low deductible insurance contracts and the disposition effect.

Obviously, the self protection problem has a continuous choice set. We therefore follow Braun and Muermann (2004) and adopt the approach of regret theoretical expected utility (RTEU) as the basis of our analysis. The RTEU approach features an arbitrary choice set and is consistent with both Sugden (1993)'s axiomatic approach and Quiggin (1994)'s Irrelevance of Statewise Dominated Alternatives (ISDA) assumption. It assumes regret is expressed as a function of the difference between the utility that would be obtained from the foregone optimal decision and the utility obtained from the actual decision:

$$\psi(x,s) = \phi(x,s) - k \cdot g\left[\phi\left(x^{opt}\left(s\right), s\right) - \phi(x,s)\right], \tag{4}$$

where x is the choice variable, s is the realization of the random state variable, ϕ is the classic von Neumann-Morgenstern utility as a function the choice and the state, $x^{opt}(s)$ is the foregone optimal action given the realized state s, g with $g > 0$,$g' > 0$, $g'' > 0$ and $g(0) = 0$ represents regret, and $k \geq 0$ stands for the intensity of regret aversion. Hence, the decision problem is written as:

$$\max_x \mathbb{E}_s\psi(x,s) = \mathbb{E}_s\left\{\phi(x,s) - k \cdot g\left[\phi\left(x^{opt}\left(s\right), s\right) - \phi(x,s)\right]\right\}, \tag{5}$$

where the subscript indicates the variable with respect to which the expectation is taken. Interestingly, without TT, the concept of regret seems almost incompatible with the self-protection problem: the decision-maker would never be able to find out the foregone optimal decision since all states within the loss (no loss) event are not distinguishable from each other. However, x^{opt} becomes revealed if TT is available in combination with the ex post observation of the exogenous risk factors. Consider again the hurricane example. Suppose there is TT, i.e. the decision-maker knows the conditional threshold effort given each potential intensity of the hurricane as well as the probability distribution of the intensity. Based on this information, he chooses effort x^* that protects his house from a hurricane up to the intensity y^*. In addition, he anticipates four potential scenarios in the future: (1) The hurricane is weaker than y^*, his house remains safe, but he would have obtained the exact same outcome had he put in some less effort. (2) The hurricane has exactly the intensity y^* and his house remains safe. (3) The hurricane is stronger than y^* and destroys his house. But he would have avoided the loss had he chosen a higher effort. (4) The hurricane is so strong that the loss is absolutely unpreventable. Any effort undertaken would be therefore completely wasted. In all four future scenarios, the decision-maker will realize ex post what he *should have done* in the past, which, except for in the second scenario, is different from what he actually did and therefore generates regret. If the future regret is anticipated ex ante, it will in turn affect the optimal effort since the decision-maker would want to minimize the expected amount of regret. Formally, the scenarios above are summarized by the following objective function:

$$\begin{aligned}\max_x V(x) =& F(0)\left[u(w) - x - k \cdot g(x-0)\right] + \int_0^x \left[u(w) - x - k \cdot g(x-t)\right] f(t)\mathrm{d}t \\ &+ \int_x^{\hat{x}} \left[u(w-L) - x - k \cdot g(u(w) - u(w-L) + x - t)\right] f(t)\mathrm{d}t \\ &+ \int_{\hat{x}}^{\infty} \left[u(w-L) - x - k \cdot g(x-0)\right] f(t)\mathrm{d}t,\end{aligned} \tag{6}$$

The first line of Equation 6 is the RTEU of situations where the actual effort exceeds the conditional threshold effort, no loss occurs and the decision-maker regrets spending too much effort. Note that the first line is decomposed into two parts due to the discontinuous distribution of the conditional threshold effort at 0. The second line stands for when the actual effort is lower than the conditional threshold effort, the loss occurs and the the decision-maker regrets spending too little effort. The third line is when the loss is unpreventable and the decision-maker regrets spending any effort at all. Note that if we eliminate g from Equation

6, it collapses to the original decision problem in our benchmark case described by Equation 1. The optimal effort of a regret-averse decision-maker x^r is therefore determined by the first order condition of Equation 6, which is obtained by applying the Leibniz rule:

$$\begin{aligned} V'(x^r) = f(x^r)\hat{x} + f(x^r)kg(\hat{x}) - \int_0^{x^r} kg'(x^r - t)f(t)\mathrm{d}t - \int_{x^r}^{\hat{x}} kg'(\hat{x} + x^r - t)f(t)\mathrm{d}t \\ - \left[1 - F(x^r)\right] kg'(x^r) - F(0)kg'(x^r) - 1 \\ = 0. \end{aligned} \tag{7}$$

By evaluating the sign of $V'(x^*)$, we can compare the optimal effort of a regret-averse decision-maker with that of an expected utility maximizer. This approach, whose detailed process is documented in the appendix, leads to the following proposition.

Proposition 4. *With TT and ex post observable risk factors, the demand for self-protection increases with regret aversion.*

Generally speaking, an increase of effort is always associated with two types of marginal benefits and two types of marginal costs. On the one hand, since the loss probability becomes further reduced, the decision-maker is both more likely to end up with higher wealth and less likely to regret letting a preventable loss occur. On the other hand, the effort itself costs more and the amount of regret increases due to the higher sunk cost. Taken together, when evaluated at x^*, the net effect of higher effort is positive due to the convexity of g. In other words, because the decision-maker is disproportionally averse to large regrets, anticipating future regret makes him willing to undertake more self-protection ex ante. TT plays an essential role in this process by revealing the conditional threshold effort, which is the crucial reference point without which a regret-averse decision maker would not be able to objectively attribute the observed event to internal (his effort) or external (the realizations of the exogenous risk factors) causes. An interesting related question is how would a regret-averse decision-maker behave without TT, which we briefly address in the next section.

5 Discussion

In the previous sections, we analyzed the impact of TT according to when *all* exogenous risk factors are observed. In reality, it may well be the case that some of the exogenous risk factors are observed ex ante, while others can only be observed ex post. In this case, we may decompose the problem into two steps. First, the ex ante observation of some, but not all risk factors corresponds to the acquisition of imperfect information, or a partial resolution of uncertainty, which leads to a Bayesian update of the probability function $p(x)$. Once the imperfect information is incorporated into the updated probability function, the rest of the problem is equivalent to the one with only ex post observable risk factors. Since imperfect information also generates non-negative value, the coexistence of ex ante and ex post observable exogenous risk factors does not change our key results.

Admittedly, in some applications, the collection of exogenous risk factors may be enormous and their interaction with the preventive effort may be highly complex. As a result of this complexity, TT may be sometimes very costly or even impossible to fully achieve. In other cases, the exogenous risk factors may never be observable, such as when a job seeker never realizes the true reason for the company to reject his application. Whenever TT or the observability of exogenous risk factors is unavailable, the ex ante removal of risk is ruled out and we are left with the question of how regret aversion can possibly affect self-protection when it is unclear how much the decision-maker should regret. One possible answer to this question is regret is irrelevant when the foregone optimal decision is unknown. However, we argue this argument is unlikely to hold since counterfactual thinking has long been documented in psychology (see Zeelenberg, 1999, for a survey). In fact, soon after the original regret theory was founded, Bell (1983) analyzed regret with unknown consequence of the foregone action when the choice set is binary, where he discusses a potential willingness to pay to avoid resolving the outcome of the alternative action. More recently, Gabillon (2018) extends the discussion to an arbitrary choice set. Both discussions are based on pre-assumed normative conditions on the decision-maker's preference. Parallel to the normative approach, phenomena such as the hindsight bias (Christensen-Szalanski and Willham, 1991), outcome bias(Baron and Hershey, 1988), or different attributional styles (Abramson et al., 1978) all suggest there is an innate tendency for people to subjectively assign reasons to past events even when they do not possess adequate information to do so. Furthermore, the same phenomena also suggest people's beliefs under insufficient knowledge may be strongly subject to behavioral biases.

However, instead of joining either the normative or the descriptive discussion, we argue that the self-protection problem has an intrinsic feature that may make both types of discussions unnecessary. In fact, this feature even makes our results in Section 4 generalizable to cases without TT. In self-protection, the consequences of any two actions (effort levels) are inherently correlated since in any state, as soon as one effort succeeds (fails) to prevent the loss, any effort larger (smaller) than it is bound to succeed (fail) as well. Hence, once the decision-maker observes the actual success/failure of his chosen effort, he can incorporate this information by performing a Bayesian update of the distribution of the conditional threshold effort. Then, if the amount of regret he experiences equals the *conditional expected regret* based on the updated distribution, his objective function will collapse to exactly our Equation 6 in the previous section. Therefore, regret may increase the ex ante self-protection effort even in the absence of TT provided that the decision-maker anticipates experiencing the conditional expected regret. This approach is in line with the concept of a stochastic reference point applied by Kőszegi and Rabin (2006) on loss aversion and by Delquié and Cillo (2006) on disappointment aversion. However, We do not deny the potentially considerable amount of cognitive effort required by this process and therefore acknowledge the importance of behavioral biases that might be relevant for this problem. For instance, the actual amount of regret may be much higher than the conditional expected regret if the decision-maker has extremely pessimistic attributional style and attributes every failure to himself and every success to luck. An extremely optimistic decision-maker, on the other hand, believes in the exact opposite and experiences much less regret. Incorporating potential biased beliefs into the analysis also requires understanding whether a decision-maker subject to biases can foresee the future biases ex ante (see the similar distinction between naïve and sophisticated present bias in O'Donoghue and Rabin, 1999). Depending on the foreseeability and the type of biases involved, TT may have very different behavioral and welfare consequences. Our aforementioned approach can be seen as a starting benchmark against which the consequences of biased beliefs may be evaluated.

6 Conclusion

In this paper, we propose the concept of technological transparency (TT) by explicitly interpreting the mechanism of self-protection as the interactive determination of the occurrence of a loss event by exogenous risk factors and the decision-maker's preventive effort. TT translates the loss probability into the joint distribution of the exogenous risk factors, which manifests itself in the distribution of the conditional threshold effort. While TT per se means knowing what determines the success of the preventive effort, its impact on the optimal effort depends crucially on when those determinants are observed by the decision-maker.

When the exogenous risk factors are observed ex ante, conditional on this observation, TT corresponds to the acquisition of perfect information and induces the most efficient effort, which prevents every preventable loss and saves all unnecessarily wasted costs. It therefore always raises the decision-maker's ex ante welfare regardless of raising or lowering the effort ex post. TT also enhances the value of technological improvement by allowing the latter to be fully (globally) exploited.

When the exogenous risk factors are observed ex post, TT allows the decision-maker to objectively attribute the (non-)occurrence of a loss to himself and to external causes. This hindsight makes the ex ante self-protection effort increase with the degree of regret aversion. In addition, TT in combination with ex post observable risk factors also serves to prevent potential biased beliefs from distorting the impact of regret aversion.

Our findings suggest that more efficient and adequate risk reduction may be enabled by scientific research uncovering exogenous risk factors and their interaction with preventive effort. In addition, the value of such scientific research may be enhanced by technological progress reducing the cost of measuring the exogenous risk factors, as well as predictive analytics helping forecast unknown values of the exogenous risk factors. Our results also have implications for the design of public education campaigns aiming to promote preventive activities. Not only should a prevention technology be made transparent by research, it needs to be made transparent in the eyes of the decision-maker. Instead of communicating the effectiveness of self-protection activities based on population average statistics, policymakers may consider tailoring the information to different subpopulation to the extend allowed by current knowledge so that each subpopulation (i.e. where the exogenous risk factors have the same realization) may choose the effort most suitable to their needs. Surprisingly, our results show that even when individuals are not yet perfectly aware of which subpopulation they belong to, simply anticipating knowing it in the future suffices to increase the current preventive effort.

References

Abramson, L. Y., Seligman, M. E., and Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of abnormal psychology*, 87(1):49.

Aumann, R. J. (1976). Agreeing to disagree. *The annals of statistics*, pages 1236–1239.

Baron, J. and Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of personality and social psychology*, 54(4):569.

Bell, D. E. (1982). Regret in decision making under uncertainty. *Operations research*, 30(5):961–981.

Bell, D. E. (1983). Risk premiums for decision regret. *Management Science*, 29(10):1156–1166.

Blackwell, D. (1951). Comparison of experiments. Technical report, HOWARD UNIVERSITY Washington United States.

Bleichrodt, H., Cillo, A., and Diecidue, E. (2010). A quantitative measurement of regret theory. *Management Science*, 56(1):161–175.

Braun, M. and Muermann, A. (2004). The impact of regret on the demand for insurance. *Journal of Risk and Insurance*, 71(4):737–767.

Briys, E. and Schlesinger, H. (1990). Risk aversion and the propensities for self-insurance and self-protection. *Southern Economic Journal*, 57(2):458–467.

Camille, N., Coricelli, G., Sallet, J., Pradat-Diehl, P., Duhamel, J.-R., and Sirigu, A. (2004). The involvement of the orbitofrontal cortex in the experience of regret. *Science*, 304(5674):1167–1170.

Christensen-Szalanski, J. J. and Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational behavior and human decision processes*, 48(1):147–168.

Courbage, C., Rey, B., and Treich, N. (2013). Prevention and precaution. In Dionne, G., editor, *Handbook of Insurance*, Huebner International Series on Risk, Insurance, and Economic Security, chapter 8. Springer.

Crocker, K. J. and Snow, A. (2000). *The Theory of Risk Classification*, pages 245–276. Springer Netherlands, Dordrecht.

Danaei, G., Ding, E. L., Mozaffarian, D., Taylor, B., Rehm, J., Murray, C. J., and Ezzati, M. (2009). The preventable causes of death in the united states: comparative risk assessment of dietary, lifestyle, and metabolic risk factors. *PLoS medicine*, 6(4):e1000058.

Delquié, P. and Cillo, A. (2006). Disappointment without prior expectation: a unifying perspective on decision under risk. *Journal of Risk and Uncertainty*, 33(3):197–215.

Dionne, G. and Eeckhoudt, L. (1985). Self-insurance, self-protection and increased risk aversion. *Economics Letters*, 17(1–2):39–42.

Ehrlich, I. and Becker, G. (1972). Market insurance, self-insurance, and self-protection. *Journal of Political Economy*, 80(4):623–648.

Fishburn, P. C. (1982). Nontransitive measurable utility. *Journal of Mathematical Psychology*, 26(1):31–67.

Friedman, M. and Savage, L. (1948). The utility analysis of choices involving risk. *Journal of Political Economy*, 56(4):279–304.

Gabillon, E. (2018). When choosing is painful: A psychological opportunity cost model. *Working Paper*.

Gollier, C. (2001). *The Economics of Risk and Time*. The MIT Press, Cambridge, MA.

Gould, J. P. (1974). Risk, stochastic preference, and the value of information. *Journal of Economic Theory*, 8(1):64 – 84.

Hilton, R. W. (1981). The determinants of information value: Synthesizing some general results. *Management Science*, 27(1):57–64.

Hoy, M. and Polborn, M. K. (2015). The value of technology improvements in games with externalities: A fresh look at offsetting behavior. *Journal of Public Economics*, 131:12–20.

Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.

Lee, C.-M. (2015). Technological advances in self-insurance and self-protection. *Economics Bulletin*, 35(3):1488–1500.

Li, L. and Peter, R. (2018). Should we do more when we know less?: Optimal risk reduction under technological uncertainty. *Working Paper*.

Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, 92(368):805–824.

Marx, J. (2002). Unraveling the causes of diabetes. *Science*.

Menegatti, M. (2009). Optimal prevention and prudence in a two-period model. *Mathematical Social Sciences*, 58(3):393–397.

Muermann, A., Mitchell, O. S., and Volkman, J. M. (2006). Regret, portfolio choice, and guarantees in defined contribution schemes. *Insurance: Mathematics and Economics*, 39(2):219–229.

Muermann, A. and Volkman Wise, J. (2006). Regret, pride, and the disposition effect.

Neumuller, S. and Rothschild, C. (2017). Financial sophistication and portfolio choice over the life cycle. *Review of Economic Dynamics*, 26:243–262.

O'Donoghue, T. and Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1):103–124.

Peter, R., Richter, A., and Steinorth, P. (2016). Yes, no, perhaps? premium risk and guaranteed renewable insurance contracts with heterogeneous incomplete private information. *Journal of Risk and Insurance*, 83(2):363–385.

Quiggin, J. (1994). Regret theory with general choice sets. *Journal of Risk and Uncertainty*, 8(2):153–165.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell system technical journal*, 27(3):379–423.

Sugden, R. (1993). An axiomatic foundation for regret theory. *Journal of Economic Theory*, 60(1):159–180.

Zeelenberg, M. (1999). Anticipated regret, expected feedback and behavioral decision making. *Journal of behavioral decision making*, 12(2):93–106.

Zeelenberg, M. and Pieters, R. (2007). A theory of regret regulation 1.0. *Journal of Consumer psychology*, 17(1):3–18.

A Appendix: mathematical proofs

A.1 Proof of Proposition 1

Since $\tilde{t} \geq 0$ by definition, $F(t) = 0$ whenever $t < 0$. If $t > 0$,

$$\begin{aligned} F(t) &= Prob(\tilde{t} \leq t) \\ &= 1 - p(t). \end{aligned}$$

Since $F(0) = 1 - p(0) > 0$, F is continuous and twice differentiable everywhere except at 0. Hence, the conditional threshold effort follows a mixed type distribution that is discrete at 0 and continuous everywhere else. It follows that the probability density function $f(t) = F'(t) = -p'(t)$ whenever $t \neq 0$. At 0, F is discontinuous and f does not exist.

A.2 Proof of Proposition 2

The decision problem given the conditional threshold effort t corresponds to the following:

$$\max_{x \geq 0} U(x;t) = u(w)\mathbb{I}\{x \geq t\} + u(w-L)\mathbb{I}\{0 \leq x < t\} - x,$$

whose solution is t whenever $t < \hat{x} = u(w) - u(w-L)$ and 0 otherwise.

A.3 Proof of Corollary 1

When $\tilde{t} \leq x^*$, the loss is prevented at the cost $\tilde{t}$ with TT, whereas it is prevented at the cost x^* without TT. This happens with the probability $F(x^*) = 1 - p(x^*)$.
When $x^* < \tilde{t} \leq \hat{x}$, the loss is prevented at the cost $\tilde{t}$ with TT, whereas it incurs despite the cost x^* being wasted without TT. This happens with the probability $F(\hat{x}) - F(x^*) = p(x^*) - p(\hat{x})$.
When $\tilde{t} > \hat{x}$, the loss is unpreventable and no effort is spent under TT, whereas the cost x^* is wasted without TT. This happens with the probability $1 - F(\hat{x}) = p(\hat{x})$.

A.4 Proof of Proposition 3

The overall value of TT is measured by the difference between the value functions with and without TT. Without TT, we have:

$$U^{bc} = U(x^*) = [1 - p(x^*)]u(w) + p(x^*)u(w-L) - x^*.$$

With TT, every preventable loss is prevented at the cost of the conditional threshold effort following the distribution specified in Proposition 1, whereas no effort is wasted upon unpreventable losses. Therefore, the expected utility before the observation of the exogenous risk factors reads:

$$\begin{aligned} U^{TT} &= F(0)u(w) + [1 - F(\hat{x})]u(w - L) + \int_0^{\hat{x}} [u(w) - t]f(t)\mathrm{d}t \\ &= [1 - p(0)]u(w) + p(\hat{x})u(w - L) - \int_0^{\hat{x}} [u(w) - t]p'(t)\mathrm{d}t \\ &= [1 - p(0)]u(w) + p(\hat{x})u(w - L) - u(w)[p(\hat{x}) - p(0)] + \hat{x}p(\hat{x}) - \int_0^{\hat{x}} p(t)\mathrm{d}t. \end{aligned}$$

Hence, the value of TT equals

$$U^{TT} - U^{bc} = x^* p(x^*) + x^* - \int_0^{\hat{x}} p(t)\mathrm{d}t.$$

It follows that the value of TT can also be represented by the area of the shaded regions in Figure 2. Obviously, if both x^* and $p(x^*)$ remains the same but $p(x)$ becomes lower for every other x, the first two terms of Equation A.4 are unchanged while the third term gets closer to zero, making the overall value of TT larger.

A.5 Proof of Proposition 4

We first show that any regret-averse decision-maker demands more self-protection than an expected utility maximizer. Inserting x^* to the left hand side of Equation 7 and applying Equation 3 yields:

$$\begin{aligned} \frac{V'(x^*)}{k} &= f(x^*)g(\hat{x}) - \int_0^{x^*} g'(x^* - t)f(t)\mathrm{d}t - \int_{x^*}^{\hat{x}} g'(\hat{x} + x^* - t)f(t)\mathrm{d}t - [1 - F(\hat{x})]\, g'(x^*) - F(0)g'(x^*) \\ &= \int_0^{x^*} g'(t)f(x^*)\mathrm{d}t + \int_{x^*}^{\hat{x}} g'(t)f(t)\mathrm{d}t - \int_0^{x^*} g'(t)f(x^* - t)\mathrm{d}t - \int_{x^*}^{\hat{x}} g'(t)f(\hat{x} + x^* - t)\mathrm{d}t \\ &\quad - [F(0) + 1 - F(\hat{x})]\, g'(x*) \end{aligned}$$

The last step follows from integration by subsitution. We may then apply the mean value theorem and obtain the following:

$$\begin{aligned} \frac{V'(x^*)}{k} &= g'(x_1) \int_0^{x^*} [f(x^*) - f(x^* - t)]\, \mathrm{d}t + g'(x_2) \int_{x^*}^{\hat{x}} [f(x^*) - f(\hat{x} + x^* - t)]\, \mathrm{d}t - [F(0) + 1 - F(\hat{x})]\, g'(x^*) \\ &> g'(x_2) \left\{ \int_0^{x^*} [f(x^*) - f(x^* - t)]\, \mathrm{d}t + \int_{x^*}^{\hat{x}} [f(x^*) - f(\hat{x} + x^* - t)]\, \mathrm{d}t \right\} \\ &= [F(0) + 1 - F(\hat{x})] \left[g'(x_2) - g'(x^*)\right] \\ &> 0 \end{aligned}$$

where $0 < x_1 < x^* < x_2 < \hat{x}$. Next, we show that an increase of regret aversion, i.e. an increase of k raises the demand for self-protection. To do this, it suffices to show $V_{xk} > 0$ due to the implicit function theorem, where V_{xk} is the cross derivative of V with respect to x and k:

$$\begin{aligned} V_{xk} &= f(x)g(\hat{x}) - \int_0^x g'(x-t)f(t)\mathrm{d}t - \int_x^{\hat{x}} g'(\hat{x}+x-t)f(t)mathrmdt - \left[1 - F(\hat{x}) + F(0)\right] g'(x) \\ &= \frac{V_x + 1 - f(x)\hat{x}}{k} \\ &= \frac{1 - f(x)\hat{x}}{k} \\ &> 0. \end{aligned}$$

The last inequality follows from $f(x) < \frac{1}{\hat{x}}$, which holds because $x > x^*$ and $f(x^*) = 1/\hat{x}$.