Understanding of the Conservatism Implicated in the Regulatory Defaults in Terms of Risk

Seung-Cheol Jang [a*]
[a] *Korea Atomic Energy Research Institute, Daedeok-Daero 989-111, Yuseong-Gu, Daejeon, 34057, Korea*
[*]*Corresponding author: scjang@kaeri.re.kr*

1. Introduction

Most individuals and agencies in the world have always endeavored to avoid undesirable risks, or at least to bring them under control. Despite these efforts, however, new risks that are highly difficult to manage continue to emerge from the use of high technologies, such as nuclear power, high-speed railway, chemicals, aircrafts, and so on. In seeking to control these risk issues, it is necessary to impose several types of regulations on those responsible for the risks, thus ensuring that they are the most effective ways to reduce risks, or to allocate limited resources to do this. Ideally, the optimal balance between a relevant measure of benefit and cost should be produced in this regulatory process. However, the question of whether or not regulatory defaults should be set conservatively has long been controversial[1]. The opponent views it as needlessly costly and irrational, and the proponent as a form of protection against possible omissions or underestimation of risks. Currently, agencies differ widely in their approaches to regulatory defaults, and the implications of these differences are not well understood as yet. For example, in the EPA risk assessment guidance for the Superfund program[2], the approved defaults for a variety of quantities are described as "90^{th}-percentile," "reasonable upper-bound," and "reasonable worst case." In the nuclear power industry, by contrast, defaults for their risk analyses have generally been set at or near the mean of the industry to determine the right priorities for the risks. It is because the adoption of conservative defaults can cause irrelevant priorities of the risk-critical components, so-called a shadow effect[3].

More importantly, regulators and regulated parties have systematically different goals or utility functions[3]. In particular, regulators have a natural incentive (and in fact often a mandate) to seek large safety margins (*e.g.*, by ensuring that risks are estimated conservatively). However, the cost of complying with regulations may be a secondary consideration for regulators. Regulated parties also have an incentive (in fact, a direct financial incentive) to ensure the safety of businesses that they own and operate, but in their case this is balanced by a competing desire to minimize their costs. Given the changes in some industries (*e.g.*, the increased competition), the urgency of a cost minimization is if anything likely to increase in the next few years. Therefore, once a regulated company has achieved a level of safety that is acceptable from a corporate point of view, it will generally have an incentive to ensure that the risks disclosed to regulators are not overestimated, in order to avoid additional burdensome regulation and the reduced operational flexibility that will be likely to result.

This paper focuses on the effects of different levels of a conservatism implicated in the regulatory defaults on the estimates of a risk. Note that we do not take any position on the merits of conservatism *per se*, but rather explore the effects of different levels of conservatism, and their implications. Understanding of the conservatism implicated in regulatory defaults in terms of a risk can help decision makers evaluate the levels of a safety likely to result from their regulatory policies.

2. Notations and General Formulation

Fortunately, the effect of conservatism implicated in regulatory defaults is a topic that is amenable to fairly rigorous mathematical analysis, using simple but plausible models of regulated party behavior. In particular, let X be the (uncertain) estimate of risk (or a risk-related parameter such as a component failure rate) that would result from a risk analysis performed using realistic parameter values and assumptions. Let us assume that the variability of X across the population of regulated parties is described by the probability function, $f_X(x)$. Furthermore, let D be the default value chosen for the same quantity. For example, if a regulated party elects to use the default rather than a realistic analysis, the same value D would be used by any regulated party in the population, regardless of its value of X. Thus, it is reasonable to assume that the risk estimate disclosed to regulators by a regulated party, Y, depends on the behaviors of regulated parties as follows.

$$Y = h(X, D) \quad (1)$$

where h is the function to represent behaviors of regulated parties. Finally, the expectation of $T\ (= Y/X)$, $E[T]$, will be adopted as a simple measure to evaluate the effect of conservatism implicated in a particular regulatory default (D) on the estimates of risk.

3. Several Measures of Conservatism in Regulatory Defaults

3.1 Measures of MGE and MGEE

First, Bier and Jang[3] assumed that the regulated party has perfect knowledge about its value of X (*e.g.*, it has already done a realistic risk analysis and is deciding whether to disclose the results to regulators). So, they suggested the risk estimate disclosed to regulators by a regulated party as follows.

$$Y = X \wedge D \quad (2)$$

where $X \wedge D$ represents minimum of both quantities, $\{X, D\}$. In other words, it means that regulated parties will disclose realistic risk estimates when they are more favorable than the approved default, and will use the default value when that is more favorable. For convenience' sake, the expectation of T defined in reference[3], will be called maximum gross effect (MGE) to differentiate from other measures suggested newly in the paper.

MGE can be obtained in closed form for an arbitrary distribution of regulated population as follows. (Refer to Appendix A)

$$MGE = \int_0^1 \frac{D}{t} \cdot f_X\left(\frac{D}{t}\right) dt + F_X(D) \quad (3)$$

where F_X is the cumulative distribution function (CDF) of X.

MGE value ranges over (0,1), and means that the risk estimate disclosed by regulated party will be on average $[1 - MGE] \times 100\%$ lower than the real risk estimate. Also, note that this degree of underestimation is an upper bound on the effect that might be observed in the real world, since the behavior of a regulated party assumes perfect gaming, *i.e.*, perfect choice of the minimum to disclose with the perfect knowledge about the value of X.

Preliminary analysis of this model has been undertaken for a wide variety of choices of the distribution $f_X(x)$, as shown in Table 1. The second column of Table 1 presents some MGE results for a few distributions. Here, some results for particular distributions such as uniform and exponential distributions were analytic results, while others for less tractable distributions were based on simulation. The results of this analysis suggest that if the default D is set equal to the expected value of the quantity of interest across the regulated population, then MGE will typically be between 0.85 and 0.96. In other words, the disclosed risk estimates will be on average 4% to 15% lower than the results of realistic risk analysis. Similar results were also obtained for other parameter values, and for the gamma and beta distributions. Bier and Jang [3] provide more discussions on the results.

More importantly, even if the estimate of the average risk is low by only about 15%, the most severe risks (*i.e.*, the largest values of X) will be underestimated by much more than this. The degree of underestimation at extreme risk, so-called maximum gross effect of extreme (MGEE), can be defined as follows.

$$MGEE(X_{(n)}) = E\left[\frac{\min(X_{(n)}, D)}{X_{(n)}}\right] \quad (4)$$

where $X_{(n)}$ denotes the maximum value among observations taken from n facilities (in increasing order), *i.e.*, realistic risk estimate at the worst site. A measure of MGEE type may be more important to regulatory matter since the degree of anticipated underestimation at the most severe risks would be much higher than that at the average risks. Considering the distribution of the largest value of X, the probability that all of n independent observations on a continuous variate are less than x is $[F_X(x)]^n$, which may be calculated approximately as $\exp(-n \cdot [1 - F_X(x)])$ by the first approximation in Taylor expansion of $\ln F_X(x)$. The probability density function (PDF) of $X_{(n)}$ is given by $n \cdot [F_X(x)]^{n-1} \cdot f_X(x)$. Thus, the expectation of $T_{(n)}$ $\left(= \frac{\min(X_{(n)}, D)}{X_{(n)}}\right)$ is obtained in a closed form as follows.

$$MGEE(T_{(n)}) = n \cdot \int_0^1 \frac{D}{t} \cdot \left[F_x\left(\frac{D}{t}\right)\right]^{n-1} \cdot f_X\left(\frac{D}{t}\right) dt + [F_X(D)]^n \quad (5)$$

In a hypothetical population of 100 nuclear power plants, the right-hand column of Table 1 shows that the risk at the worst plant can be underestimated by an order of magnitude. All of the results were based on simulation due to less tractable distributions.

Table 1. Underestimation of Risks Using Mean Value Defaults* (MGE, MGEE)

Distribution	MGE	MGEE
Exponential	0.85**	0.20 ± 0.01
Weibull (shape parameter 2)	0.88 ± 0.02	0.40 ± 0.01
Weibull (shape parameter 3)	0.90 ± 0.02	0.51 ± 0.01
Weibull (shape parameter 5)	0.92 ± 0.01	0.66 ± 0.01
Uniform (lower bound=0)	0.85**	0.505 ± 0.001
Lognormal (EF=3)	0.88 ± 0.02	0.24 ± 0.01
Lognormal (EF=10)	0.90 ± 0.02	0.10 ± 0.01
Lognormal (EF=30)	0.92 ± 0.01	0.06 ± 0.01

* Error bounds for simulation results are ± 2σ

** Mean values for analytic results (Refer to App. A)

3.2 Measures of MPE and MPEE

MGE[3] measures gross average on the degree of underestimation due to defaults. According to circumstances, however, regulators may have an attribute to be more concerned about only the degree of pure underestimation of regulated risks (*i.e.*, only the case of $X > D$), because they have a natural tendency to seek large safety margins as mentioned before. Moreover, if they have to set a new regulatory default, they may concern about maximum pure underestimation on the risks disclosed to them by a regulated party in the future. Thus, the risk estimate disclosed to regulators by a regulated party will be simply defined as $Y = D$, given $X > D$. MGE can be no longer appropriate for reflecting such tendency of regulators, since it is the gross averaged measure of underestimation over the whole range of X. Measure of MGE has also a property that the more left-skewed is the distribution of a regulated population (*e.g.*, a lognormal with a long tail), the less degree of underestimation may result.

Considering the diverse concerns of regulators on their regulatory problems, another measure, so-called maximum pure effect (MPE) can be suggested, as follows.

$$MPE = E(T|T<1) = E\left(\frac{D}{X}|X>D\right) = \int_0^1 \frac{D}{t}\cdot f_X\left(\frac{D}{t}\right)dt \quad (6)$$

Note that MPE is defined as a conditional expectation and corresponds to the first term in the right hand side of equation (3) which is related to MGE. In other words, it means the pure effect of underestimation due to the default specified by regulators.

Preliminary simulation analyses of this model have been undertaken for lognormal distributions with equal median (0.001) but different error factors (*e.g.*, 3, 10, 30, respectively), as shown in Table 2. The results of simulation analyses show that the magnitude of the maximum pure effect of underestimation is in reverse order, compared with ones of MGE. In other words, the more left-skewed with a long tail rightwards is the distribution of a regulated population, the more degree of underestimation may result. Figure 1 shows the difference between MPE and MGE for the lognormal distributions with different error factors in detail. It can be regarded as the inevitable gap between the regulator and regulated party, which can occur frequently in the process of a risk-informed decision making.

Similar to MPE, the conditional expectation of order statistics of extreme risk, so-called Maximum Pure Effect of Extreme (MPEE) can also be defined as the first term in the right-hand of equation (5) which is related to MGEE.

$$MPEE(T_{(n)}) = n\cdot\int_0^1 \frac{D}{t}\cdot\left[F_x\left(\frac{D}{t}\right)\right]^{n-1}\cdot f_X\left(\frac{D}{t}\right)dt \quad (9)$$

Table 2. Underestimation of Risks Using Mean Value Defaults* (MPE, MGE)

Distribution	MPE	MGE
Lognormal (EF=3)	0.66	0.87
Lognormal (EF=10)	0.50	0.88
Lognormal (EF=30)	0.47	0.92
Lognormal (EF=100)	0.44	0.96

*Error bounds for simulation results are ± 2σ

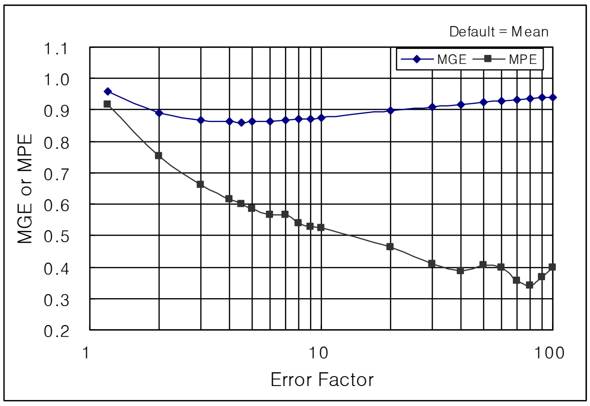

Fig. 1. Comparison between MGE and MPE for Lognormal Distribution (Median=0.001)

3. Conclusions

The desirability of conservatism in regulatory risk analyses has long been controversial. It is seen by some as needlessly costly and irrational, and by others as a form of a protection against possible omissions or underestimation of risks.

The intractability of this debate may arise in part because it views conservatism in isolation, rather than as one element of an overall regulatory system. This paper focuses on the effects of different levels of a conservatism implicated in the regulatory defaults on the estimates of a risk. Note that we do not take any position on the merits of conservatism *per se*, but rather explore the effects of different levels of conservatism, and their implications. Understanding of the conservatism implicated in regulatory defaults in terms of a risk can help decision makers evaluate the levels of a safety likely to result from their regulatory policies.

ACKNOWLEDGEMENTS

This work was supported by the Nuclear Research & Development Program of the National Research Foundation of Korea grant, funded by Korean government, Ministry of Science, ICT & Future Planning.

REFERENCES

[1] National Research Council (1994), *Science and Judgment in Risk Assessment*, National Academy Press, Washington, D.C.

[2] U.S. Environmental Protection Agency (1991), *Risk Assessment Guidance for Superfund, Volume I: Human Health Evaluation Manual, Supplemental Guidance*, "Standard Default Exposure Factors," Interim Final, Office of Solid Waste and Emergency Response, OSWER Directive 9285.6-03, Washington, D.C.

[3] Bier, V. M., and S. C. Jang (1999), "Perspective: Defaults and Incentives in Risk-Informed Regulation", Human and Ecological Risk Assessment 5, Number 4, pp.635-644.

Appendix A. *Proof of MGE and Analytic Solutions*

By definition, the quantity of interest, X , is a positive variate. Then, the domain of $T(=Y/X)$ becomes $0 \le T \le 1$. Because of the mass probability at $T=1$ (i.e., $X \le D$), only cumulative density function (CDF) of T , $G_T(t)$, may be derived as described below.

$$G_T(t)=\begin{cases}0 & , T<0\\ 1-F_X\left(\frac{D}{t}\right) & , 0\le T<1\\ 1 & , T\ge 1\end{cases} \quad \text{(A1)}$$

Reimann-Stieltjes integration leads from equation (A1) to the expectation of an arbitrary function of T as follows.

$$E[h(T)]=\int_0^1 h(t)\cdot dG_T(t)+h(1)\cdot \Delta G_T(1) \quad \text{(A2)}$$

where $\Delta G_T(1)$ means the mass probability at $T=1$, and corresponds to $F_X(D)$. Substituting T for $h(T)$ in equation (A2), the expectation of T can be obtained in a closed form as shown equation (3).

As an illustration, the analytical results obtained for uniform and exponential distributions in Table 1 can be derived from equations (3). If the quantity of interest across the regulated population follows a uniform distribution over $[a,b]$, the expectation of T are given as follows.

$$E(T)=\frac{D}{b-a}\cdot\left\{1+\ln\left(\frac{b}{D}\right)\right\} \quad \text{(A3)}$$

Note that $D/b \le T \le 1$. If D is set to the mean of X , i.e.,$(a+b)/2$, and $a=0$, then the expectation of T become 0.8466, regardless of the value of b . It means that the risk estimate disclosed by regulated party is maximum 15% lower than the realistic risk estimate. In other words, this presents the maximum effect that may be underestimated by the regulatory default of the mean value.

For a quantity of interest, X , which follows an exponential distribution with parameter λ , expectation of T may be approximately calculated using a Taylor series expansion as follows.

$$E(T)=1-e^{-\xi}-\xi\cdot\left\{A+\ln\xi+\sum_{n=1}^{\infty}\frac{(-\xi)^n}{n\cdot n!}\right\} \quad \text{(A4)}$$

where $\xi=\lambda D$ and A stands for Euler's constant (0.57721...). If regulatory default is equal to the mean, $1/\lambda$, then mean of T are approximately given by 0.8515, as shown in Table 1.