Incorporating Ex-Ante Risk in Evaluating Public R&D Programs: A Counterfactual Analysis of the Korean Case

So Young Kim*

Abstract

R&D is inherently an uncertain endeavor, yet now more than ever those performing R&D with public funding are called upon to clarify the utility of their research. Calls for public accountability are mounting with the increase in constraints on government budgets due to the recent worldwide economic recession, in response to which both policymakers and researchers pay much more attention to rigorously assessing publicly funded R&D. A key issue complicating R&D evaluation in these circumstances is how to adequately account for the nature and degree of risk involved in a given R&D program or project. This study deliberates on certain issues involving the measurement of ex-ante risk in public R&D evaluation: (i) information asymmetry between R&D sponsors and performers, (ii) ambiguity in the measurement of returns in both prospective and retrospective evaluation, and (iii) the dilemma between measurement error and omitted variable bias for empirical estimation of R&D performance. The study then presents an analysis of hypothetical evaluation results that apply risk-relevant weights to the annual evaluation outcomes of South Korea's national R&D programs funded during 2006~2012. In this counterfactual re-evaluation of public R&D program performance, high-risk R&D programs turn out to receive higher evaluation than non-high-risk programs. The current study suggests that R&D evaluation ignoring ex-ante risk is not only conceptually invalid since R&D activities are intrinsically uncertain endeavors, but unfair as R&D performers are asked to be accountable for the results that were in fact out of their reach.

Keywords
evaluation, high-risk high-return, honorable failure, information asymmetry, R&D, uncertainty

* Associate Professor, Graduate School of Science and Technology Policy, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, soyoungkim@kaist.ac.kr

1. INTRODUCTION

Research and development (R&D) activities contain many layers of uncertainty. Much of this uncertainty comes from a long temporal gap between input and output in the R&D process. In addition, a complex interplay of a large number of actors in our contemporary R&D environment also adds increasing difficulty in predicting a final output or ensuring an intended outcome of R&D.

As it does in any other sphere of public policy, uncertainty raises many challenges for sponsors, planners, and implementers of publicly funded R&D. First and foremost, it is a reason behind information asymmetry between different R&D actors, often generating perverse incentives to exploit imperfect or incomplete information of other actors. Uncertainty also challenges R&D actors by testing their level of trust. The long gestation times typical of scientific research and technological development, a vital aspect of R&D, are a result of repeated interactions over uncertain futures and is unsustainable without trust and respect among researchers as well as R&D sponsors.

Yet in the R&D context, uncertainty is more of a double-edged sword than a curse. It is because of uncertainty that remarkable inventions enter the markets and groundbreaking discoveries revolutionize society. An unexpectedly positive R&D outcome is the holy grail of R&D investors, public or private. While those investing significant capital into R&D should expect normal returns from their investment, an expectation of unusually high return in exchange for larger risk is built into their motivation. The latter premise has recently served a strong rationale for various public R&D funding schemes such as high-risk/high-return, high-risk/high-gain, frontier, or transformative research programs.

This study reviews various ways in which the term "risk" is appropriated in the R&D context as well as recent attempts to incorporate "ex-ante risk" into R&D evaluation. It then discusses certain issues that complicate measuring and integrating ex-ante risk into evaluation in publicly funded R&D programs and projects, followed by a counterfactual analysis of the evaluation results of publicly-funded R&D programs in South Korea over the past seven years (2006~2012). This analysis applies different weights to the evaluation of R&D programs selected experts identified as riskier but yielding higher returns than other programs. The final section concludes with a summary and discussion of the study's implications.

2. CONCEPTUAL TURNABOUT OF THE TERM

When first introduced into the English vocabulary in the seventeenth century, "risk" meant a peril that could compromise a long voyage (*risque* derived from Latin *riscum*), which was viewed as an objective danger attributed to an act of God (Lupton 1999, p.5). The modern usage of the term underwent two transformations in meaning. First, risk was no longer attributed to supernatural forces but signified an outcome of *human* action. Second, risk began to take a distinctively normative tone by referring to the *unanticipated* consequences of such action. As a deviation (an unexpected outcome) from a norm (an intended effect), it came to take on a strongly negative connotation.

In the early 1900s, noted economist Frank Knight (1921) came to neutralize the negative connotations of the term in his treatise on risk and uncertainty. Criticizing the ambiguities of the two notions, he called for a clear distinction between them stating that "a measurable uncertainty, or 'risk' proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all. We ... accordingly restrict the term 'uncertainty' to cases of the non-quantitative type" (p.121, emphasis added). In this conceptual turnaround, he offered an operational definition of risk as the potential effect of an event or an activity multiplied by its likelihood. Defined as such, risk has become an indispensable part of business management and finance. In the valuation of a market investment portfolio, for instance, risk and return form an integrated assessment of the value of the portfolio, with risk typically measured as the standard deviation of the average return of a given investment.

In the R&D community, the discourse on risk has evolved along two contrasting lines. In the research on and practice of technology assessment, risk is associated with potential dangers and hazards resulting from the economic and social applications of emerging technologies. In his now classic analysis of nuclear accidents and other technological disasters, Charles Perrow (1984/1999) characterizes high-risk technology as being prone to system failures, involving unanticipated interactions of multiple disruptions that affect more than one subsystem. Technology assessment in this context is, broadly speaking, risk assessment applied to such high-risk technology. As such, the literature on technology assessment and the related fields such as technology foresight and futurology is replete with analyses identifying, describing, communicating, and managing risk from novel technological advances enabled by R&D. In this community of researchers and practitioners, risk has negative connotations and is almost indistinguishable from hazards.

In marked contrast, R&D sponsors have lately employed the term of risk in a more positive tone. The US National Science Foundation (NSF) was one of the first funding agencies that deliberately used "high-risk research" to encourage creative research proposals spanning multiple disciplines. It has introduced multiple funding programs for high-risk research since 2007 including High-Risk Research in Biological Anthropology and Archeology (HRRBAA), Integrated NSF Support Promoting Interdisciplinary Research and Education (INSPIRE), Creative Research Awards for Transformative Interdisciplinary Ventures (CREATIV), and Early-concept Grants for Exploratory Research (EAGER).[1]

1. While some of these programs do not employ the term high risk per se, universities and the news media such as the Chronicles of Higher Education cover the awardees as winning under a new NSF high-risk funding scheme. The EAGER program explicitly targets high-risk projects, as stated in its description: "The EAGER funding mechanism may be used to support exploratory work in its early stages on untested, but potentially transformative, research ideas or approaches. This work may be considered especially 'high risk-high payoff' in the sense that it, for example, involves radically different approaches, applies new expertise, or engages novel disciplinary or interdisciplinary perspectives" (National Science Foundation 2013).

In Europe, the well-known initiative for the creation of the European Research Area aims to achieve excellence in frontier research by encouraging researchers to be more risk-taking. The European Commission defines frontier research as an "intrinsically risky endeavor that involves the pursuit of questions without regard for established disciplinary boundaries or national borders" (Leijten, Roseboom & Hofer, 2010, p.21).

A non-Western country these calls for high-risk research have struck a chord with is South Korea. Initial pushes for high-risk research began in universities and other research organizations,[2] the momentum of which was picked up by the national government in the form of a concrete plan to reward high-risk research in national R&D programs. In a 2012 proposal for high-risk research promotion, the Korean government defined high-risk R&D as follows: "while the risk of failure is high given the capacities of domestic R&D performers, high-risk R&D programs will generate world-class scholarly achievement or technological breakthroughs that can benefit a large number of industries or create new ones" (Korean Government 2012, p.2).

In short, different factions of the R&D community understand the term "risk" in diametrically opposite fashions. For those mostly associated with technology assessment and foresight, risk is something that should be avoided or minimized in the process of developing and deploying new technologies. For those mostly involved in R&D funding, risk is something we should actively encourage researchers to take up in order to create new possibilities in inherently uncertain endeavors to advance knowledge and make technical breakthroughs.[3] In the following two sections, we focus on risk in relation to R&D evaluation, first reviewing the recent attempts to incorporate ex-ante risk into R&D evaluation, and then examining potential problems that may generate tensions in evaluating high-risk R&D work.

3. RISK IN R&D EVALUATION

Noting the apparent lack of attention to risk in the retrospective evaluation of R&D projects, Linquiti (2012) recently called upon the R&D community to develop a scheme that incorporates ex-ante risk into ex-post R&D evaluation, suggesting the use of methodologies from financial risk

2. The High-Risk High-Return Program (HRHR) of the Korea Advanced Institute of Science and Technology (KAIST) was launched in 2008 to support research projects that are highly promising but hard to obtain external funding for due to high risk of failure. The Korea Academy of Science and Technology (KAST) held a special roundtable on KAIST's HRHRP to explore its implications for the existing practices of R&D funding, evaluation, and national research culture (Korea Academy of Science and Technology 2008).

3 In an informal survey taken at the author's institution in fall 2012, undergraduate students majoring in science and technology policy were asked to select words or expressions they associate the term "risk" with. Interestingly, those taking courses on science communication (which treats risk perception and communication as a course topic) overwhelmingly suggested words such as "hazards, jeopardy, danger, or failure," whereas most of the students taking courses on R&D management associated risk with "adventure, challenge, quality, novelty," or "Nobel Prize."

assessment. He first identifies two types of risk in R&D investment, as in Link and Scott (2011). *Technical risk is the risk* of failure due to technical challenges or poor project execution, and is therefore inherent to an R&D project. *Market risk reflects* uncertainties about realizing the commercialization potential of an R&D output whereupon R&D agents are faced not only with uncertainty intrinsic to the nature of research but also to uncertainty in the external environment, which, while typified by any market, encompasses the policies or processes that could compromise the initial strategies and plans of the agents.

According to Linquiti, prospective R&D evaluation naturally takes ex-ante risk into account as most sponsors of R&D projects clearly have in mind how likely the selected projects are to succeed. (After all, public and private R&D investors alike have to come up with good reasons to shareholders or taxpayers for their investment, and high probability of success is one of the best ways to persuade them.) It is actually retrospective R&D evaluation that fails to pay due attention to ex-ante risk. In his example reproduced in a modified form in Table 1, two multi-year R&D Projects X and Y are all the same except the probability of success. With the higher probability of success, Project Y is expected to have a higher "expected" internal rate of return (IRR), which would be factored into the project selection process. Once the projects are completed, the "observed" IRR for both projects would depend only on whether the project has actually succeeded or not.[4] In other words, ex-ante risk does not play any role in this ex-post evaluation, and we would have no idea whether X or Y resulted in a successful outcome.

TABLE 1. Ex-ante Risk in R&D Evaluation

Prospective Evaluation		X	Y
	Annual Cost	1	1
	Capital Cost	35	35
	Annual Benefit	10	10
	Probability of Success	0.1	0.9
	"Expected" IRR	8.1	21.6
"Expected" Annual Net Benefit			
	Years 1-5	-10	-10
	Year 6	-35	-35
	Years 7-26	1	9
Retrospective Evaluation		Success	Failure
	Observed IRR	22.1	~-100.0
Annual Net Benefit			
	Years 1-5	-10	-10
	Year 6	-35	0
	Years 7-26	10	0

(Cost and benefit figures in million dollars) Modified from Linquiti (2012)

4. As is well known in finance literature, the IRR is the discount rate that makes the net present value (NPV) of a proposed project zero. The IRR is a popular metric to use when comparing alternative project options as it offers a simple rule for project choice, which is to undertake the project if the IRR is greater than the discount rate. Yet it can be misleading when assessing a project with large up-front benefits and long-term costs (or vice versa). Since R&D projects typically have large start-up costs and long rather than short-term benefits, a simple comparison based on the IRR can result in a rejection of an R&D project that may yield significant benefit over a long period.

Another attempt to incorporate risk into R&D evaluation is found in recent studies of technology valuation. Chen et al. (2009) in their assessment of Taiwanese renewable technology portfolios proposes indicators for two metrics (importance and risk) to assess R&D projects. Technology risk is then measured by five indicators which reflect the levels of technology and potential for commercialization: technology position (gap between domestic and global technology levels), manufacturing capability (whether the quantity and quality of technology output meets international standards), industrial supply chains (upstream-downstream coordination), market entry barriers, and lead time for commercialization. In a similar vein, Hsieh (2013) develops a framework for strategic patent valuation using two dimensions (risk and benefit) where R&D risk is classified into finer categories such as technological risk, managerial risk, market acceptance risk, production risk, and other cost-related risks. What is notable in these studies is that although their attempts to account for the degree of risk in technology valuation are fairly new and innovative, their assessments are made on the basis of choice about energy technologies or patents that would be selected or maintained by policymakers, and therefore consider risk only in the context of prospective evaluation.

4. ISSUES IN MEASURING EX-ANTE R&D EVALUATION

While it makes intuitive sense that ex-ante risk should be taken into account in retrospective R&D evaluation as it does in prospective evaluation, there are certain issues that complicate such endeavors.

4.1. Information Asymmetry

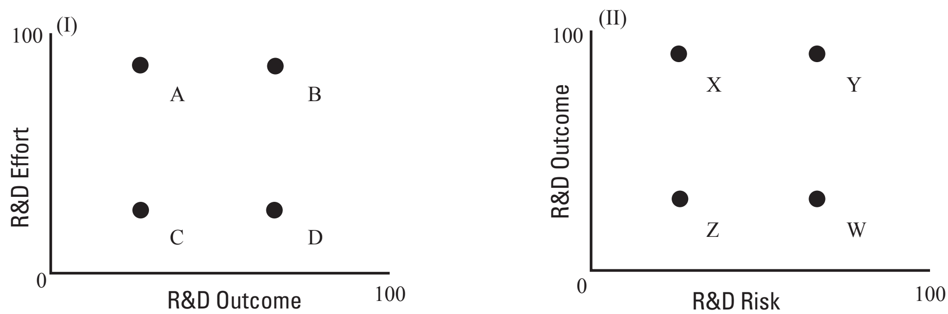
R&D evaluators conducting ex-post assessment will inevitably suffer from asymmetric information between them and the R&D performers, as it is difficult to determine accurately whether a given project entailed great or small risk and whether researchers took risks or did not.[5]

To clarify the problem of information asymmetry in the context of R&D evaluation, consider four R&D projects under ex-post evaluation shown in Figure 1. The first panel of this figure places the projects on two dimensions: whether a researcher or a team of researchers conducting a government-sponsored R&D project has put in much effort, and whether the outcome was successful. Researchers have better information on both dimensions as they are by definition experts on their research areas. In contrast, R&D evaluators are poorly informed as to how much the researchers have strived to achieve the purported outcome, although they may be on par with the researchers

5. Information asymmetry between policymakers and academics/researchers is a source of tension and contradiction in science and technology policy. For instance, Guston (2000) applies the principal-agent theory to analyze how the integrity and productivity of research can be assured in the face of conflicting demands from science and politics. This theory hinges upon information asymmetry between the parties to a contract such as doctor-patient, lawyer-client, etc. With one party (i.e. an agent) having more or better expertise in the work involved, there emerges an incentive to exploit the lack of knowledge of the other party (i.e. a principal).

in assessing the outcome (given the fact that many evaluators are also recruited from the relevant research community). In other words, it is relatively easy for the evaluators to distinguish Projects B and D from Projects A and C, but hard to do so for A/B vs. C/D. In particular, a key question for R&D evaluators concerns the assessment of A vs. D (note that it would be straightforward to compare B and C, as both projects score consistently high or low). Would or should R&D evaluators give higher points for Project A whose researchers tried hard towards the goal but did not obtain a great outcome, or for Project D that resulted in a better outcome despite smaller R&D effort? [6]

FIGURE 1. Information Asymmetry in Measuring Ex-ante Risk



A-D and X-W indicate hypothetical R&D projects

R&D Risk: technical and market risk entailed in R&D projects

R&D Outcome: R&D performance including direct results and long-term impact

This asymmetry of information as to the level of research effort is compounded by yet another layer of asymmetric information on the degree of risk that researchers are exposed to and/or taking up, as illustrated in the second panel of the figure. Again, in ex-post evaluation that considers R&D projects upon their completion, R&D outcome is comparatively easy to determine. Less clear is how much risk a given project actually entails, about which people performing R&D must have a better estimate than those observing their effort. If R&D sponsors were risk-averse, researchers would have an incentive to present their research as safe (low-risk) and higher-return, such as Project X. In contrast, if R&D sponsors were encouraging risk-taking as has been the case in recent funding schemes mentioned earlier, researchers would want their work to be viewed as high-risk and high-return, such as Project Y.

6. This is also a matter of equity and fairness. The way the current question is posed would lead many to agree that high-effort researchers of Project A should not be disadvantaged just because of a poor outcome. The same argument may in fact apply the other way round. It would not be fair to give less regard for low-effort researchers of Project D, as they might have made relatively little effort not because they were lazy or uninterested but because they were competent enough to complete the project without much effort. In other words, there exists a problem of evaluating absolute performance vs. achievement relative to competency in addition to information asymmetry. See Stone (2002) for an excellent discussion of a tradeoff between efficiency and equity in many *cul-de-sac* policy situations.

Another related question is how evaluators should evaluate risk-taking on the part of researchers when the latter produces poor outcomes (that is, between Z and W). This is a question that goes right into the heart of the recent calls for acknowledging "honorable failure" in R&D evaluation in order to encourage researchers to challenge more daunting problems. Central to this concept is the idea that if researchers have ventured into high-risk research yet failed to produce significant output, they should not be disadvantaged in their performance evaluation. Given the asymmetry of information between R&D funders/evaluators and R&D performers as to the nature and degree of risk involved, however, it is a challenge as to how the former can distinguish researchers who did their best but "honorably" failed in the face of high risk from those who "dishonorably" failed because they did not push themselves to the best.

4.2. High Benefit vs. High Payoff

When speaking of high-risk, high-return projects, there appears to be some confusion over what "returns" actually mean. Are we encouraging researchers to take up more risk because we want higher returns or higher payoffs? To illustrate this point, let us consider three projects presented in Table 2. Project X is expected to have the least cost and the smallest benefit but is the most likely to succeed. Project Z has the highest expected costs and the largest expected benefit but has the highest probability of failure. Project Y is in between X and Z in terms of cost, benefit, and risk. In a prospective evaluation given different levels of risk, the expected payoff is ordered by magnitude as X > Y > Z. What is unclear about the recent high-risk and high-return funding schemes is whether they are trying to maximize the benefit (in which case projects should be evaluated in the order of Z > Y > X) or the payoff (in which case the evaluation result would be reversed).

TABLE 2. High Risk, High Return, High Payoff

	Prospective Evaluation				Retrospective Evaluation			
					A		B	
	Expected Cost	Expected Benefit	Chance of Success	Expected Payoff	Outcome	Observed Payoff	Outcome	Observed Payoff
X	1	5	0.7	3.2	Success	4	Failure	-1
Y	5	10	0.5	2.5	Success	15	Failure	-5
Z	10	30	0.3	2	Success	25	Failure	-10

(Cost and benefit figures in million dollars)
For simplicity, the realized cost/benefit in retrospective evaluation is assumed to be the same as the expected cost/benefit.

In retrospective evaluation, we would also have completely opposite evaluation results depending on how these projects were realized. For simplicity, let us assume that the actual resulting costs and benefits are the same as the expected ones and that all projects turn out to be successful. Then the actual payoff (now simply the difference between the realized benefit and the realized cost) would be highest for Z, which promised the highest benefit. If all projects are unsuccessful however, ex-post evaluation would score them in completely reverse order.

4.3. Measurement Error vs. Omitted Variable Bias

There is also a less theoretical and more practical issue when empirically measuring ex-ante risk in R&D evaluation. The value of an R&D project is a function of many factors. Suppose we are evaluating the performance of a certain R&D project, say *Y*, with *k* variables as follows:

$$P(y)=f(x_1, x_2, x_3, \ldots, x_k)$$

If we want to estimate the above equation with a regression model, we will plug in a specific measure for each variable such that $x_j = X_j + \varepsilon_j$ where ε_j represents measurement error. Since it is much more complicated to measure the degree of risk compared to other variables relevant to R&D performance (such as full-time researchers, R&D equipments, funding, etc.), ε_j will be typically larger for the risk variable than most other variables. Given such an imprecise measure, we may forgo measuring risk. This decision, however, invites a problem of omitted variable bias, if risk is indeed one of the key drivers of the project outcome. As is well known in introductory statistics, linear regression estimators in the presence of omitted variables are no longer unbiased and consistent.[7] In short, in empirical estimation of R&D performance we face an inherent dilemma in our choice of incorporation of ex-ante risk into R&D evaluation: should we tolerate inaccuracy in the measurement of R&D risk or run the risk of omitted variable bias?

The next section presents an analysis of the national R&D programs funded by the South Korean government since the introduction of a formal R&D evaluation system in 2006. Rather than offering a full-fledged analysis, this is a more general survey of a method of incorporating ex-ante risk into ex-post evaluation by selected experts giving extra weights to programs viewed as high-risk.

5. COUNTERFACTUAL R&D EVALUATION INCOPORATING EX-ANTE RISK

Despite the long history of strategically appropriated R&D for economic growth being one of the world's largest R&D expenditures,[8] it was not until the mid-2000s that the groundwork for a systematic scheme of R&D evaluation was laid. With the introduction of the Performance Evaluation and Performance Management of National R&D Act ("R&D Performance Act") in December 2005, the South Korean government began to annually evaluate government-funded R&D programs. In the span of seven years from 2006 to 2012, a total of 785 R&D programs were evaluated. Due to

7. Note that measurement error in the dependent variable still gives unbiased estimates of the parameters, but measurement error in an explanatory variable leads to not only biased but inconsistent estimates even if sample size increases indefinitely (Gujarati and Porter 2009, p.484).
8. As of 2012, South Korea's gross R&D expenditure (GERD) amounts to 4.03% of GDP, which ranks the nation at the second on the R&D intensity (only after Israel). Even in absolute amount of R&D spending, South Korea ranks sixth after much larger countries (US, Japan, China, Germany, and France).

government reorganizations during this period after the presidential election, specific agencies in charge of evaluation as well as actual schemes of evaluation have varied by year. However, two common grade scales (A-B-C-D or 0~100/0~200) continued to be used over the different evaluation periods.[9]

Using actual evaluation results for nationally funded R&D programs, this study undertakes a counterfactual analysis of the results that might be derived if high-risk programs were given extra weights.[10] The evaluation results are available from the *Statistical Summary of the National R&D Programs* published by the Bureau of Performance Evaluation of the Korean National Science and Technology Commission. For statistical analysis, actual evaluation results recorded in this summary were converted to numeric data (for example, A-B-C-D grades are coded as 4-3-2-1).

The most important part of this analysis is to classify R&D programs in terms of the degree of risk. For this task, we conducted a survey in February 2013 of ninety-eight prominent South Korean scientists and engineers. Selected on the basis of winning prestigious awards in science and engineering, these experts were asked to select R&D programs they would consider as high-risk in their own area of expertise.[11] Although it would have been ideal to have them classify the risk they associated with a specific R&D program into "market risk" and "technical risk" as in the literature presented earlier, the survey in order to increase the response rate asked them to assess only the degree of risk. Given that most of these experts are university researchers relatively distant from market situations, we may interpret the risk they had in mind as technical risk rather than market risk.

Table 3 shows the results of hypothetical evaluation using the assessments of R&D risk derived from the above expert survey, which is an essentially counterfactual endeavor as R&D risk is retrofit into the actual evaluation outcome. Not surprisingly, the number of R&D programs considered high-risk (HR) is smaller than that of non-HR programs. It is also noteworthy that for four out of seven years, HR programs received lower grades than non-HR programs (as shown in the fourth

9.Another common feature of this annual R&D performance evaluation is a distinction between self- and higher-level evaluations. Self-evaluation is evaluation conducted internally within an R&D program whereas higher-level valuation reviews evaluation results submitted by the researchers. This study relies on self-evaluation results, although it would be more interesting and relevant to use higher-level evaluation results. This choice is because there were notable changes in agencies and implementation methods for higher-level evaluation (for example, it took the form of meta-evaluation in some years and re-evaluation in other years).

10 .The initial plan was to conduct this R&D evaluation analysis at both the program and project levels. However, project-level evaluation analysis turned out to be more daunting than expected, given the considerably large number of projects that makes it virtually impossible to conduct an expert survey (explained shortly) of counterfactual assessment of risk.

11. Reflecting both the size of funding and the research community, different numbers of experts were selected in nine areas (ten for mathematical sciences, twelve for physics, eighteen for biological sciences and medicine, eight for earth sciences, eleven for chemistry/chemical engineering/materials science, eight for mechanical engineering, nine for electrical engineering, twelve for information communication technology, and ten for civil engineering and transportation). The initial version of the questionnaire asked the experts to rank the programs according to degree of risk, which resulted in poor responses. The final version was streamlined to induce a maximum response rate, having respondents only pick a program they considered high-risk. A total of thirty-four experts replied (34.7% response rate), with certain fields such as mathematics returning more responses than others. More details about the survey design are available upon request.

row where HR programs were unweighted). However, multiplying the evaluation grades of the HR programs with 1%, 2.5%, and the standard deviation of each year's evaluation grades gives HR programs consistently higher grades than non-HR programs, with most HR and non-HR differentials also being statistically significant on the difference of means tests. Table 4 shows the same analysis applied to evaluation scores (note that the 2006 performance evaluation did not use the scores). The results are largely similar to those for the evaluations using the grade scale. The current finding implies that if ex-post assessment of R&D performance took into account the presence of risk, R&D programs taking up higher risk would have been more favorably evaluated.

TABLE 3. Hypothetical R&D Evaluation Incorporating Ex-ante Risk (Evaluation Grades)

R&D Program	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
# of High-Risk (HR) Programs	39	32	39	12	14	22	14
# of NonHR Programs	128	122	125	58	56	65	59
Mean Evaluation Grades	3.401	2.403	3.018	2.871	3.543	2.851	2.717
HR-NonHR Differential with HR Unweighted	-0.022	-0.272	0.615***	0.055	0.125	-0.895	-0.683
HR-NonHR Differential with HR Weighted by 1%	0.114	-0.184	0.755***	0.171	0.271*	-0.808	-0.595
HR-NonHR Differential with HR Weighted by 2.5%	0.317***	-0.053	0.964***	0.346	0.489***	-0.677	-0.462
HR-NonHR Differential with HR Weighted by SD	0.646***	0.447***	1.425***	0.795	0.755***	0.099	0.305

*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Source: Statistical Summary of the National R&D Programs 2012 (Performance Evaluation Bureau, National Science & Technology Commission of Korea)
Note: (1) The mean evaluation grades are the means of the self-evaluation grades (A=4, B=3, C=2, D=1) of all R &D programs of the respective year. (2) The HR-NonHR differentials were t-tested. (3) The last row results were obtained from weighting HR programs with the standard deviation of the respective year's evaluation grades.

TABLE 4. Hypothetical R&D Evaluation Incorporating Ex-ante Risk (Evaluation Scores)

R&D Program	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Mean Evaluation Scores	117.055	91.727	90.239	90.214	84.230	82.725
HR-NonHR Differential with HR Unweighted	-4.720	4.512	0.587	1.777	-9.890	-9.556
HR-NonHR Differential with HR Weighted by 2%	-2.454	6.416***	2.402*	3.650*	-8.361	-8.042
HR-NonHR Differential with HR Weighted by 5%	0.945	9.271***	5.123***	6.359***	-6.067	-5.771
HR-NonHR Differential with HR Weighted by SD	16.459***	10.859***	6.316***	8.267***	1.317	1.035

*: $p < 0.10$, **: $p < 0.05$, ***: $p < 0.01$

Source: Statistical Summary of the National R&D Programs 2012 (Performance Evaluation Bureau, National Science & Technology Commission of Korea)
Note: (1) See Table 3 for the number of HR and NonHR programs. (2) The mean evaluation scores are the means of the self-evaluation scores of all R&D programs of the respective year. (3) The HR-NonHR differentials were t-tested. (4) The score results are weighted by different percentages from those for the grade results due to different scales (note also that score evaluation was on the scale of 0~100 for all years except for 2007 when evaluation was conduced on the 0~200 scale). (5) The last row results were obtained from weighting HR programs with the standard deviation of the respective year's evaluation scores. (6) No score evaluation was conducted in 2006.

One caveat for the current exercise of counterfactual evaluation is that completely opposite results would have arisen had they been done on a different perception of risk. In the current analysis, most R&D experts interpreted high-risk as something benign (and closely associated with high returns), and thus positive weights were given to HR programs. If a more negative aspect of the term were emphasized, negative weights would have been applied to HR programs, resulting in very different evaluation results. Although not attempted in this analysis, it would also be interesting and significant to find out how such a counterfactual evaluation incorporating ex-ante R&D risk may yield different results by different types of R&D risk. For instance, we may expect systematic differences between R&D programs that entail large technical risk but are relatively straightforward to commercialize (i.e. low on market risk) and those with relatively small technical risk but with great difficulty in commercialization and marketing. Further research is needed in this direction.

6. DISCUSSION

While R&D inherently entails uncertainties, both sponsors and performers of R&D programs and projects, particularly those funded by taxpayer dollars, are called upon to turn in clear and confident benefits of R&D investment. With increasing constraints on government budgets due to the worldwide economic recession, the mounting pressure makes government S&T policymakers concentrate limited research funding on R&D efforts worth supporting in light of market and public values (Link and Scott 2011). In other words, we are witnessing a trend referred to as “the new social contract” where scientists and engineers are increasingly called upon to explain their research outcomes and demonstrate the utility thereof in return for public funding for their research (Jasanoff 2007).

This rising demand for accountability of publicly-funded R&D is making both policymakers and researchers pay more attention to ensuring proper returns on public R&D investment, which in turn generates a drive towards a more rigorous assessment of the performance of government-funded R&D.

A key issue complicating R&D evaluation in this context is how we can adequately account for the nature and degree of risk involved in a given R&D program or project. Risk is viewed in certain segments of the R&D community as a hazard posed by new discoveries or inventions, whereas in other sections of the community it is something researchers should more actively take up in order to achieve technological breakthroughs or scientificadvance our understanding of nature.

This study deliberates on certain issues that arise when measuring and incorporating ex-ante risk into public R&D evaluation: information asymmetry between the sponsors/evaluators and the performers of R&D, ambiguity in the measurement of returns as benefits or payoffs in both prospective and retrospective evaluation, and a dilemma between measurement error and omitted variable bias for empirical estimation of R&D performance.

Also presented in this study is a brief analysis of hypothetical evaluation results applying extra weights to the actual evaluation outcomes of South Korea’s national R&D programs conducted in the past seven years. Based upon the expert survey where the nation’s prominent scientists and engineers were asked to identify high-risk R&D programs, the past R&D evaluation results were recalibrated using different weights. In this counterfactual re-evaluation of the public R&D program performance, high-risk R&D programs turned out to receive higher evaluation (whether measured on a scale of grades or scores) than non-high-risk programs.

In order to encourage government-funded researchers to take on more challenging problems, it is crucial to design a funding scheme that rewards those who take up greater risk of failure. To this end, R&D policymakers need to better understand complexities in identifying high-risk high-return R&D projects as well as measuring the level of R&D efforts on the part of researchers in the presence of risk and uncertainty. One of the most significant implications of the current study is that R&D evaluation ignoring ex-ante risk of any R&D effort is not only conceptually invalid since R&D activities are intrinsically uncertain endeavors, but unfair as R&D performers are asked to be accountable for results that were in fact out of their reach. It is tremendously difficult to design an evaluation scheme in the presence of all the challenges in ex-ante risk measurement (in the face of a reality where R&D evaluation as the basis of funding allocation is a great tug-of-war among all societal stakeholders in public R&D spending), yet clarifying exactly what contradictions or dilemmas the funders and sponsors of public R&D face when they do try to appraise R&D performance will be a crucial step towards a more constructive R&D evaluation system that is a more efficient use of taxpayers' money.

REFERENCES

American Academy of Arts and Sciences. (2008). Advancing research in science and engineering: Investing in early-career scientists and high risk, high reward research. *Report for the Committee on Alternative Modes for the Federal Funding of Science*. Washington, D.C.: AAAS

Arrow, K., & Lind, R. (1970). Uncertainty and the evaluation of public investment decisions. *American Economic Review*, 60, 364–78.

Chen, T., Yu, O., Hsu, G., Hsu, F., & Sung, W. (2009). Renewable energy technology portfolio planning with scenario analysis: a case study for Taiwan. *Energy Policy,* 37(8), 2900-2906.

Guston, D. H. (2000). *Between politics and science: Assuring the integrity and productivity of research*. New York: Cambridge University Press.

Gujarati, D. N., & Porter, D.C. (2009). *Basic econometrics* (5th ed). New York: McGraw-Hill.

Hsieh, C. (2013). Patent value assessment and commercialization strategy. *Technological Forecasting and Social Change*, *80,* 307-319.

Jasanoff, S. (2007). New social contract for science. *In Designs on nature: science and democracy in Europe and the United States*. Princeton: Princeton University Press.

Klein, M. (1997). The risk premium for evaluating public projects. *Oxford Review of Economic Policy, 13*, 29-42.

Knight, F. H. (1921). *Risk, uncertainty, and profit*. Boston: Houghton Mifflin.

Korea Academy of Science and Technology (KAST). (2008). How to promote high-risk high return R&D? *Report for the 54th KAST Roundtable*. Seoul: KAST.

Korean Government. (2010). Proposals for promoting risk-taking in national R&D programs. *Briefing for the office of prime minister, national science & technology commission, and nine other ministries.* Seoul.

Leijten, J., Roseboom, H., & Hofer, R. (2010). More frontier research for Europe: A venture approach for funding high risk – high gain research. *Report for the European Commission Director of General Research*. Brussels: EU Commission Director-General for Research and Innovation.

Linquiti, P. (2012). The importance of integrating risk in retrospective evaluations of R&D. *Research Evaluation, 21,* 152–165.

Link, A. &Scott, J. (2011). *Public goods, public gains: Calculating the social benefits of public R&D*. New York: Oxford University Press.

Lupton, D. (1999). *Risk.* New York: Routledge.

National Science and Technology Commission of Korea (NSTC). (2012). *Statistical summary of the national R&D programs*. Bureau of Performance Evaluation. Seoul: NSTC.

National Science Board. (2007). Enhancing support of transformative research at the National Science Foundation. Arlington, VA: National Science Board.

National Science Foundation. (2013). Introduction to transformative research. Retrieved from http://www.nsf.gov/about/transformative_research/index.jsp.

Perrow, C. (1984/1999). *Normal accidents: living with high-risk technologies*. Princeton: Princeton University Press.

Stone, D. (2002). *Policy paradox: The art of political decision making*. New York: W. W. Norton.