Management Science

Invariant Probabilistic Sensitivity Analysis

Manel Baucells, Emanuele Borgonovo

MANAGEMENT SCIENCE
Vol. 59, No. 11, November 2013, pp. 2536–2549
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2013.1719

Invariant Probabilistic Sensitivity Analysis

Manel Baucells
RAND Corporation, Santa Monica, California 90401; and Department of Economics and Business, Universitat Pompeu Fabra, 08018 Barcelona, Spain, baucells@rand.org

Emanuele Borgonovo
Department of Decision Sciences and ELEUSI, Bocconi University, 20136 Milan, Italy, emanuele.borgonovo@unibocconi.it

In evaluating opportunities, investors wish to identify key sources of uncertainty. We propose a new way to measure how sensitive model outputs are to each probabilistic input (e.g., revenues, growth, idiosyncratic risk parameters). We base our approach on measuring the distance between cumulative distributions (risk profiles) using a metric that is invariant to monotonic transformations. Thus, the sensitivity measure will not vary by alternative specifications of the utility function over the output. To measure separation, we propose using either Kuiper's metric or Kolmogorov–Smirnov's metric. We illustrate the advantages of our proposed sensitivity measure by comparing it with others, most notably, the contribution-to-variance measures. Our measure can be obtained as a by-product of a Monte Carlo simulation. We illustrate our approach in several examples, focusing on investment analysis situations.

Key words: probabilistic sensitivity; investment valuation; risk analysis; decision analysis; scale invariance
History: Received January 20, 2012; accepted December 21, 2012, by Peter Wakker, decision analysis. Published online in *Articles in Advance* May 23, 2013.

1. Introduction

As part of the evaluation of business ventures, investments, mergers, and industrial projects, best practices recommend building a financial model of projected cash flows (Benninga 2008). Such a model allows the decision maker to gain insights into the investment economics. Typically, the model is a single scenario in which inputs are set at their average or most likely value. Suppose that, after a preliminary evaluation, the decision maker obtains a positive net present value (NPV). Realizing that several of the model inputs are not known with certainty, the decision maker may assign probability distributions and run a Monte Carlo simulation (MCS; see Hazen and Huang 2006).[1] The resulting cumulative distribution function (CDF) of the output is called the risk profile. When the output is an NPV, the analyst may be interested in the expected NPV, the probability of negative NPV, the value at risk, the conditional value at risk, and the standard deviation. The user may also perform sensitivity analysis and quantify the contribution of each assumption to overall uncertainty. This setup is common in quantitative analyses.

Several sensitivity analysis methods have been proposed over the years. The method that requires the least information is the tornado diagram (Howard 1983), which uses the range endpoints of each input variable and ignores probabilistic information. Variance-based methods, based on the second moments of the distributions, are more informative (Wagner 1995). Javelin diagrams and measures based on the expected value of perfect information (Felli and Hazen 1998) use all the probabilistic information but require additional information (e.g., a full specification of other available alternatives or of utilities). The contribution-to-variance method is the most popular.

Our goal is to introduce a new class of sensitivity measures that possess advantageous properties. The key idea is to measure how much an input influences the output by averaging the distance between the unconditional and the conditional (on that input) distributions of the output. We repeat these calculations for each of n inputs. Numerous metrics are available to measure the distance between risk profiles. We focus on the few that are scale invariant and, among these, propose either the Kuiper (1960) metric or the Kolmogorov–Smirnov metric. These sensitivity measures are denoted by the n-dimensional vectors

[1] The cumulative distribution function of the output is analytically available under restrictive conditions (e.g., multivariate normal). If the relationship between inputs and outputs is linear, then the expected output is obtained from setting the inputs at their expected values in the static model. In general, the relationship between inputs and output may not be linear (capacity and inventory constraints), or the average input may not be a meaningful entry (as with discrete distributions representing variables such as approval). In this case, MCS or other integration methods are required to obtain the distribution and expected value of the output.

β^{Ku} (Kuiper) and β^{KS} (Kolmogorov–Smirnov). β^{Ku} and β^{KS} are well defined for any given joint distribution and functional relationship between inputs and output (e.g., even if the second moments of some inputs are infinite). Both measures possess five distinctive properties: they (1) are scale invariant, (2) use all the probabilistic input–output information, (3) do not require additional information, (4) impose a low computational burden (being calculated from the sample generated by MCS), and (5) have an intuitive probabilistic interpretation.

Scale invariance, also called invariance to monotonic transformations, means that given any monotonic function u, the sensitivity of $u(Y)$ to any given input is independent of u. Baucells and Sarin (2007) discuss three methods to evaluate cash flows: discounting certainty equivalents, discounting the utility of cash flows, and taking the utility of the discounted cash flows. They show that the only method consistent with a preference for the early arrival of cash and the monotonicity of the utility function is the use of the utility of NPV, $u(NPV)$. By scale invariance, the sensitivity measure obtained in a risk-neutral setting remains after any utility assignment. Scale invariance is desirable in a firm's context, where it is often difficult to identify the decision maker's utility (the utility function of a chief executive officer and each of the shareholders may differ).

The probabilistic interpretation of β_i^{Ku} is an upper bound on the *increase* in the probability of success or failure produced by the input i; that for β_i^{KS} is an upper bound on the *change* in probability of success produced by the input i.

Whereas β^{Ku} and β^{KS} are defined for generic models, we focus on applying them in NPV analysis. We obtain an analytical expression for the conditional and unconditional NPV distributions for the case in which the random inputs are the cash flows, relaxing the classical independence assumption (Grubbström 1967). We illustrate the method by several test cases and apply it to a case study in the pharmaceutical sector. Our conclusions and insights do not depend on the choice of β^{Ku} versus β^{KS}. We favor and adopt β^{Ku} for its ability to weight equally all portions of a distribution function. We compare β^{Ku} with contribution to variance. In particular, we present three situations in which contribution to variance provides meaningless answers but β^{Ku} does not.

2. Comparison of Probabilistic Sensitivity Measures

Consider the dependence of a deterministic model output on n uncertain exogenous variables:

$$y(\theta)\colon\ \Omega \to \mathbb{R} \tag{1}$$

with $\theta \in \Omega \subseteq \mathbb{R}^n$. Here, Ω is the set of possible values that the inputs can assume; we use the terms "exogenous variables" and "inputs" interchangeably to denote the vector θ. The input–output mapping $y(\theta)$ may be known analytically (e.g., a formula cell in a spreadsheet model) or as the output of a computer code (e.g., a linear program).

As is customary in the literature, $\boldsymbol{\Theta}$ denotes a random vector and θ one of its realizations. The associated probability space is denoted by $(\Omega, \mathscr{B}(\Omega), \mathbb{P}_{\boldsymbol{\Theta}})$. We let $Y = y(\boldsymbol{\Theta})$ be the corresponding random model output. The distributions of $\boldsymbol{\Theta}$ and Y are denoted by $F_{\boldsymbol{\Theta}}(\theta) = \mathbb{P}_{\boldsymbol{\Theta}}(\bigcap_{i=1}^n [\boldsymbol{\Theta}_i \le \theta_i])$ and $F_Y(y) = \mathbb{P}_Y(Y \le y)$; $f_{\boldsymbol{\Theta}}(\theta)$ and $f_Y(y)$ denote their corresponding densities, if they exist. In business applications, $F_Y(y)$ is called a risk profile.

Given this setup, one may want to quantify the importance of each input $\boldsymbol{\Theta}_i$, $i = 1, \ldots, n$, on Y. Many sensitivity measures have been proposed. Table 1 offers a comparison of the main sensitivity measures with respect to five properties.

Table 1 **Comparison of Sensitivity Analysis Methods**

Property	Tornado diagrams	Javelin diagrams	Correlation based	Variance based	EVPI based	β^{Ku} and β^{KS}
Is global	No	Yes	Yes	Yes	Yes	Yes
Handles nonlinearities well	Yes	Yes	No	Yes	Yes	Yes
Handles dependencies well	No	No	No	No	Yes	Yes
Information required is $Y(\boldsymbol{\Theta})$	No (less)	No (more)	Yes	Yes	No (more)	Yes
Is scale invariant	No	No	No	No	No	Yes

Local measures address the response of the model output around a reference value of the exogenous variables, say, θ^0. The simplest approach is to monitor the values of Y as each input takes a low and a high value while keeping other inputs at the reference value. The result is two numbers, one low and one high, for each input. The method is known as, and represented by, a tornado diagram (Howard 1988). Tornado diagrams are widely used and produce much insight given their simplicity (Clemen 1997). Tornado diagrams are imprecise because they reduce probabilistic information to ranges and ignore probabilistic dependencies.

Hazen and Huang (2006) demonstrate that in the presence of correlations, the sensitivity method must be *global*—that is, account for probabilistic dependencies. Felli and Hazen (2004) introduce javelin diagrams, a generalization of tornado diagrams in the presence of uncertainty. They present an example in which the insights derived from a javelin diagram contradict those derived from a tornado diagram. A drawback of javelin diagrams, however, is their need for additional information (e.g., a full specification of other available alternatives or of utilities).

Other global sensitivity analysis methods can be categorized into four groups: correlation based, contribution-to-variance based, expected value of perfect information (EVPI) based, and density based. Correlation-based methods were historically developed first (Saltelli and Marivoet 1990) and are included as postanalysis options in professional MCS software programs such as Crystal Ball and @Risk. These methods measure the influence of $\mathbf{\Theta}_i$ on Y by the strength of their correlation. Correlation can be defined by either the Pearson correlation coefficient, $\text{cov}(Y, \mathbf{\Theta}_i)/(\sigma_i \sigma_Y)$, or the standardized regression coefficient, $b_i\sigma_i/\sigma_Y$, where b_i is the regression coefficient of $\mathbf{\Theta}_i$ in $Y = b_0 + \sum_{s=1}^{n} b_s \cdot \theta_s$. Their performance, however, rapidly deteriorates in the presence of nonlinearities and interactions, e.g., multiplications, between exogenous variables (Campolongo and Saltelli 1997).

The contribution-to-variance measure has been proposed as a more robust global sensitivity method. It is defined as

$$\eta_i^2 := \frac{\mathbb{V}\{\mathbb{E}[Y \mid \mathbf{\Theta}_i]\}}{\mathbb{V}[Y]} = 1 - \frac{\mathbb{E}\{\mathbb{V}[Y \mid \mathbf{\Theta}_i]\}}{\mathbb{V}[Y]}. \tag{2}$$

In Equation (2), η_i^2 is the expected reduction in the variance of Y, given that the exogenous variable $\mathbf{\Theta}_i$ is fixed. It was originally called the correlation ratio in Pearson (1905) and was introduced as a sensitivity measure in operational risk analysis by Iman and Hora (1990). One can also define the total order sensitivity measure:

$$\eta_i^T := \frac{\mathbb{E}\{\mathbb{V}[Y \mid \mathbf{\Theta}_{\sim i}]\}}{\mathbb{V}[Y]} = 1 - \frac{\mathbb{V}\{\mathbb{E}[Y \mid \mathbf{\Theta}_{\sim i}]\}}{\mathbb{V}[Y]}, \tag{3}$$

where $\mathbf{\Theta}_{\sim i}$ denotes the vector of all factors but $\mathbf{\Theta}_i$ (Wagner 1995). If the random inputs are independent, then η_i^2 and η_i^T provide a measure of functional dependence (Owen 2003); η_i^2 becomes the normalized individual contribution of $\mathbf{\Theta}_i$ to $\mathbb{V}[Y]$, and η_i^T is the fraction of $\mathbb{V}[Y]$ associated with all contributions of $\mathbf{\Theta}_i$ individually plus all $\mathbf{\Theta}_i$'s interactions with the remaining groups of exogenous variables.

When exogenous variables are dependent, the one-to-one correspondence between η_i^2 and η_i^T and model structure is lost (Bedford 1998, Oakley and O'Hagan 2004), although the interpretation in terms of variance reduction is retained (Wagner 1995).

Variance-based measures are sometimes subject to misinterpretation. Analysts use variance as a summary of uncertainty, concluding that the exogenous variable that most reduces variance is the one that most reduces uncertainty (Oakley and O'Hagan 2004, p. 753). This agreement holds under the restriction that the distribution of Y is fully characterized by its first two moments, as it is in the case of normal distributions. Variance contribution does not account for skewness or differences in higher moments (Cox 2008).

EVPI is a central concept in decision analysis (Howard 1966, Pratt et al. 1995). $EVPI_i$ represents the expected utility added to the decision problem by the possibility of knowing $\mathbf{\Theta}_i$ before the decision is made. EVPI is the most appropriate measure of *decision sensitivity* (Felli and Hazen 1998). Under a quadratic utility assumption, $EVPI_i$ and η_i^2 are equivalent sensitivity measures (Oakley 2009). The use of EVPI as a sensitivity measure might not always be applicable. For example, suppose that the NPV of an investment is positive for any realization of the model inputs. Then, $EVPI$ is zero for all inputs because the alternative "to invest" is always preferred. Nevertheless, the analyst would still wish to know which inputs most influence the output.

Finally, we consider density-based methods. They seek to compare two states of knowledge: before and after knowing that $\mathbf{\Theta}_i = \theta_i$. The decision maker's belief about Y changes from the unconditional density $f_Y(y)$ to the conditional density $f_{Y \mid \mathbf{\Theta}_i = \theta_i}(y)$. The effect of getting to know $\mathbf{\Theta}_i$ can then be quantified by measuring the discrepancy between $f_Y(y)$ and $f_{Y \mid \mathbf{\Theta}_i = \theta_i}(y)$, averaged over all possible realizations of $\mathbf{\Theta}_i$. Discrepancy can be measured in multiple ways, e.g., by using the Kullack–Leibler divergence measure from information theory (Park and Ahn 1994) or the statistical theory of density separation (Chun et al. 2000, Borgonovo 2007). In addition to measuring input sensitivity, distribution separation plays a fundamental role in probability elicitation (Jose et al. 2008, 2009).

The approach that we propose is to compare cumulative distribution functions rather than densities. We do so for two reasons. First, cumulative distribution functions are defined for all distributions, even if a distribution does not admit a density. Second, cumulative distributions allow us to relate the sensitivity measure to the probabilities of exceeding or falling short of a target (Bordley and Li Calzi 2000). In summary, we measure sensitivity based on the distance between the unconditional risk profile, $F_Y(y)$, and the conditional risk profile, $F_{Y \mid \mathbf{\Theta}_i = \theta_i}(y)$.

3. Distribution-Based Scale Invariant Measures

3.1. Motivation

We begin with two examples that show the shortcomings of contribution-to-variance measures.

Example 1. Let $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ be independently distributed, with $\mathbf{\Theta}_1$ uniform on $[-1, 1]$ and $\mathbf{\Theta}_2 \sim \text{Beta}(10, 10)$. Let

$$Y = \mathbf{\Theta}_1 \mathbf{\Theta}_2 + 1.$$

We check that $\mathbb{E}[Y] = 1$ and $\mathbb{V}[Y] = 6.3$. Intuitively, we would guess that both inputs are relevant and that $\mathbf{\Theta}_2$ is more important than $\mathbf{\Theta}_1$. Surprisingly, we have $\eta_1^2 = \eta_2^2 = 0$ and $\eta_1^T = \eta_2^T = 1$. That η_1^2 and η_2^2 are both zero is unsatisfactory because both $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ have an effect on Y and on its distribution. That η_1^T and η_2^T are equal to each other is also unsatisfactory because the distributions of $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ are markedly different. This negative result holds for any function of a multiplicative form whose component functions $g_i(\theta_i)$ have null expectation (see Appendix A for all proofs).

Lemma 1. *Let the output take the form*

$$y = \prod_{i=1}^{n} g_i(\theta_i) + c.$$

Assume the inputs are independent random variables and that the model is square integrable; $\mathbb{V}[g_i(\mathbf{\Theta}_i)] < \infty$, $i = 1, \ldots, n$. *If* $\mathbb{E}[g_i(\mathbf{\Theta}_i)] = 0$, $i = 1, \ldots, n$, *then for all* i, $\eta_i^2 = 0$ *and* $\eta_i^T = 1$.

Example 2. A common task in the actuarial sciences is assessing risk associated with a portfolio of n assets, with $Y = \sum_{i=1}^{n} \mathbf{\Theta}_i$. For its properties, the Pareto distribution ($F(\theta_i) = 1 - \xi_i/\theta_i^{\alpha_i}$, if $\theta_i \geq \xi_i$) may be adopted to model asset randomness (Ramsay 2008). The analyst might be interested in knowing how the different assets contribute to uncertainty in Y. If the shape parameter, α_i, is less than 2 for one or more of the assets, then Y is not square integrable, and all variance-based sensitivity measures, η_i^2 and η_i^T, $i = 1, \ldots, n$, are undefined. In this case, variance-based sensitivity measures provide no insight into which asset contributes more to portfolio variability.

3.2. The β^d Measures

What drives the previous results is that contribution to variance considers only the first two moments of the distribution. Our approach uses more information. Rather than comparing unconditional and conditional variances, we propose comparing unconditional and conditional distributions. We quantify the distance between F_Y and $F_{Y\mid \mathbf{\Theta}_i=\theta_i}$ using a probability metric, $d\{\cdot, \cdot\}$. The distance needs to be quantified for all possible values of θ_i. Taking expectations over $\mathbf{\Theta}_i$ makes the separation unconditional.

Definition 1. Given a probability metric between distributions, $d\{\cdot, \cdot\}$, we define *the importance measure of* $\mathbf{\Theta}_i$ *with respect to* Y as

$$\beta_i^d := \mathbb{E}[d\{F_Y, F_{Y\mid \mathbf{\Theta}_i=\theta_i}\}]. \tag{4}$$

One sensible property of this approach is that if Y is independent of $\mathbf{\Theta}_i$, then $\beta_i^d = 0$.[2]

There is a large variety of choices for the metric d. Possible choices include the Kolmogorov–Smirnov (KS), Kuiper (Ku), and Cramér–von Mises metrics and their generalizations (Anderson and Darling 1952). We are interested in probability metrics that are invariant to monotonic transformations and that are easy to interpret.

3.3. Metric Choice

We consider first Kuiper's (1960) metric, given by

$$d^{Ku}\{F_Y, F_{Y\mid \mathbf{\Theta}_i=\theta_i}\} = \underbrace{\sup_y\{F_Y(y) - F_{Y\mid \mathbf{\Theta}_i=\theta_i}(y)\}}_{\Delta P^s} + \underbrace{\sup_y\{F_{Y\mid \mathbf{\Theta}_i=\theta_i}(y) - F_Y(y)\}}_{\Delta P^f}. \tag{5}$$

Clearly, $0 \leq \Delta P^s + \Delta P^f \leq 1$. We provide a probabilistic interpretation of Kuiper's distance in terms of increases in probability of success or failure due to $\mathbf{\Theta}_i = \theta_i$. Let $\hat{y}$ be some success threshold. That is, if $Y > \hat{y}$, the decision maker gets a reward (e.g., promotion, prestige) or the firm has enough extra profit for reinvesting in research and development projects insuring increased competitiveness. The probability of interest is $P(Y > \hat{y}) = 1 - F_Y(\hat{y})$. If we knew that $\mathbf{\Theta}_i = \theta_i$, then this probability would become $P(Y > \hat{y} \mid \mathbf{\Theta}_i = \theta_i) = 1 - F_{Y\mid \mathbf{\Theta}_i=\theta_i}(\hat{y})$. The difference between the new (conditional) and original (unconditional) probability of success is $F_Y(\hat{y}) - F_{Y\mid \mathbf{\Theta}_i=\theta_i}(\hat{y})$. Thus, $\sup_y\{F_Y(y) - F_{Y\mid \mathbf{\Theta}_i=\theta_i}(y)\}$ is the *highest possible increase in the probability of success* upon conditioning on θ_i, independent of the location of the threshold of success, $\hat{y}$. We denote this term by ΔP^s.

Similarly, let $\hat{y}$ be some threshold of failure (e.g., if $Y \leq \hat{y}$, then the decision maker is fired or the firm has to abandon some market). The probability of interest is $P(Y \leq \hat{y}) = F_Y(\hat{y})$. If we were to know that $\mathbf{\Theta}_i = \theta_i$, then this probability becomes $P_{Y\mid \mathbf{\Theta}_i=\theta_i}(Y \leq \hat{y}) = F_{Y\mid \mathbf{\Theta}_i=\theta_i}(\hat{y})$. The difference between the conditional and the unconditional failure probability is $F_{Y\mid \mathbf{\Theta}_i=\theta_i}(\hat{y}) - F_Y(\hat{y})$. Thus, $\sup_y\{F_{Y\mid \mathbf{\Theta}_i=\theta_i}(y) - F_Y(y)\}$ is the *highest possible increase in the probability of failure* upon conditioning on θ_i, independent of the location of the threshold of failure, $\hat{y}$. We denote this term by ΔP^f.

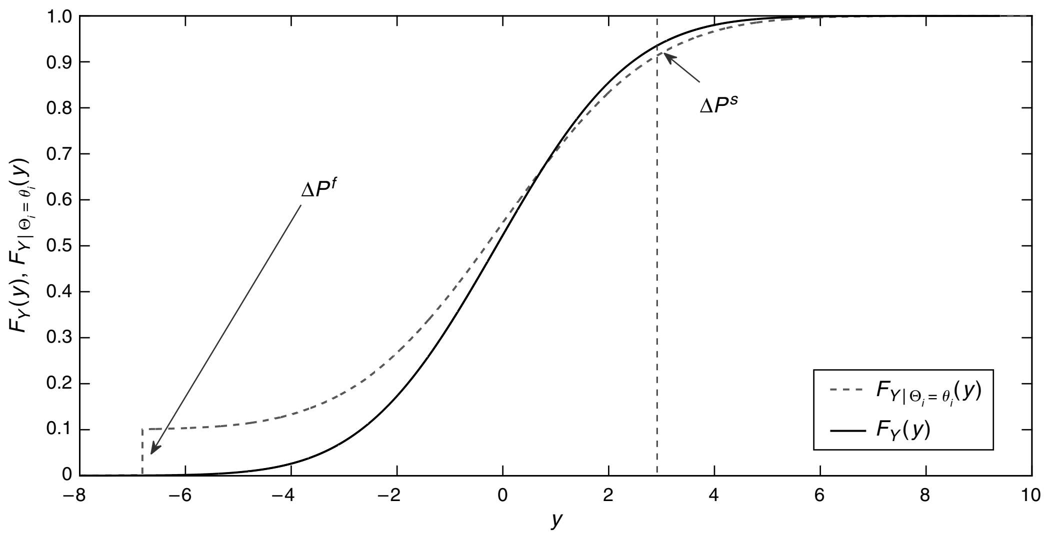
We present a visualization of Kuiper's (1960) metric in Figure 1.

Because of the laws of probability,

$$P(\mathit{Success} \text{ or } \mathit{Failure}) \leq P(\mathit{Success}) + P(\mathit{Failure}),$$

[2] Consider the joint distribution $F(Y, \mathbf{\Theta}_i)$. By definition, if Y is independent of $\mathbf{\Theta}_i$, then $F_{Y,\mathbf{\Theta}_i}(y, \theta_i) = F_Y(y) \cdot F_i(\theta_i)$, from which $F(Y \mid \mathbf{\Theta}_i) = F_Y(y)$ for all values of $\mathbf{\Theta}_i$. Hence, $d\{F_Y, F_{Y\mid \mathbf{\Theta}_i=\theta_i}\} = d\{F_Y, F_Y\} = 0 \;\forall\, \theta_i \Longrightarrow \beta_i^d = 0$.

Figure 1 **Visualization of Kuiper's Metric**



we have that $\Delta P^s + \Delta P^f$ is the highest possible increase in the probability of success or failure.[3] By taking the expectation of Equation (5), β_i^{Ku} measures the expected maximum *increase in the probability of success or failure* due to θ_i. This, we believe, is a sensible measure of sensitivity.

A second possibility is the Kolmogorov–Smirnov metric, given by

$$d^{KS}\{F_Y, F_{Y|\boldsymbol{\Theta}_i=\theta_i}\} = \sup_y |F_Y(y) - F_{Y|\boldsymbol{\Theta}_i=\theta_i}(y)| = \max\{\Delta P^s, \Delta P^f\}. \quad (6)$$

d^{KS} can be interpreted as the maximum change, positive or negative, in the probability of success upon conditioning on θ_i. By taking the expectation of Equation (6), β_i^{KS} measures the expected maximum *change in the probability of success* due to θ_i.[4] Clearly, β^{Ku} and β^{KS} are normalized measures; i.e., $\beta_i^d \in [0, 1]$, $i = 1, \ldots, n$.

Noting that $d^{Ku} = \Delta P^s + \Delta P^f = \max\{\Delta P^s, \Delta P^f\} + \min\{\Delta P^s, \Delta P^f\}$, we have

$$d^{Ku} = d^{KS} + \min\{\Delta P^s, \Delta P^f\},$$

which clearly implies $d^{Ku} \geq d^{KS}$. The equality $d^{Ku} = d^{KS}$ holds if $\min\{\Delta P^s, \Delta P^f\} = 0$ because one of the distributions, e.g., $F_{Y|\boldsymbol{\Theta}_i=\theta_i}(y)$, stochastically dominates the other, e.g., $F_Y(y)$.

Kuiper (1960) introduced the metric bearing his name to solve a problem associated with the KS metric for goodness-of-fit tests on variables distributed on a circumference rather than on the real line. Kuiper's modification allows the statistic to become independent of the starting point (Watson 1961). This ability of the Kuiper metric to "put all percentiles on an equal footing" (Crnkovic and Drachman 1996, p. 140) has made it popular in value-at-risk studies in finance (Crnkovic and Drachman 1996, Lopez and Saidenberg 2000, Berkowitz 2001).

Statisticians have emphasized that the KS metric "exhibits poor sensitivity to deviations from the hypothesized distribution that occur in the tail" (Mason and Schuenemeyer 1983, p. 933). Specifications of weighting functions that make the KS metric equally sensitive at all points or place emphasis on the tail of a distribution are discussed in Anderson and Darling (1952). Mason and Schuenemeyer (1983) propose a weighted combination of members of the Birnbaum–Orlicz family of metrics (Deza and Deza 2009). Their metric is summarized by the expression

$$d^{MS}\{F_Y, F_{Y|\boldsymbol{\Theta}_i=\theta_i}\} = \sup_y\{w[F_Y(y), F_{Y|\boldsymbol{\Theta}_i=\theta_i}(y)] \cdot |F_Y(y) - F_{Y|\boldsymbol{\Theta}_i=\theta_i}(y)|\}, \quad (7)$$

where $w[F_Y(y), F_{Y|\boldsymbol{\Theta}_i=\theta_i}(y)]$ is a weighting function depending on both the distribution chosen and the relevance one wishes to assign to different portions of the distribution. As Mason and Schuenemeyer (1983, p. 944) state, however, "We have not found a uniformly good weight function, but we have shown theoretically that there do exist weight functions ... that make particular weighted K-S tests consistent with respect to local deviations that may occur in the tails or the middle of the distribution." Crnkovic and Drachman (1996) propose the weighting function $\ln(F_Y(y)[1 - F_Y(y)])$ to make the Kuiper metric more

[3] Success and failure may not be mutually exclusive because the threshold of failure may be higher than that of success.

[4] Bordley and Li Calzi (2000) show that a decision maker selecting the alternative that maximizes $P(Y > \hat{y})$, where $\hat{y}$ is a random target, is consistent with the axioms of Savage (1954).

sensitive to tail deviations. Thus, an analyst has available a wide choice of metrics, and she can select the one that is most appropriate to her goals. Yet introducing ad hoc weighting functions would cause the metrics to lose symmetry and make them analytically less tractable. For our work, we prefer to retain the simplicity of β^{Ku} and β^{KS}, with a special focus on β^{Ku} because of its desirable properties.

To illustrate, we calculate the values of β^{Ku} and β^{KS} as applied to Examples 1 and 2.[5]

Example 1 (Continued). The results are $\beta^{Ku} = (0.28, 0.38)$ and $\beta^{KS} = (0.14, 0.19)$. Note that $\beta^{Ku} = 2 \cdot \beta^{KS}$ because of the symmetry in the distributions. These values indicate that both $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ are relevant, with $\mathbf{\Theta}_2$ being more influential than $\mathbf{\Theta}_1$—an intuitive result.

Example 2 (Continued). Even if some inputs are not square integrable, one can always rank assets using β^{Ku} and β^{KS}. For instance, let $n = 2$, $\alpha_1 = 1.3$, and $\alpha_2 = 1.5$. Then $\beta^{Ku} = (0.46, 0.39)$ and $\beta^{KS} = (0.39, 0.34)$, and η^2 is undefined.

4. Scale Invariance

Let $F_Y(\cdot)$ and $G_Y(\cdot)$ be the CDFs associated with distributions $\mathbb{P}$ and $\mathbb{Q}$, respectively. A metric d is invariant to monotonic transformations if

$$d\{F_{u(Y)}(u), G_{u(Y)}(u)\} = d\{F_Y(y), G_Y(y)\}.$$

Both d^{Ku} and d^{KS} are invariant to monotonic transformations.[6] Intuitively, monotonic transformations deform the horizontal axis of the (Y, F_Y) plane. Because β^{Ku} and β^{KS} are defined in the vertical axis, they are unaltered by such deformations.

Proposition 1. *If a metric d is invariant to monotonic transformations, then the associated importance measure is invariant to monotonic transformations; i.e.,*

$$\beta_i^d(Y) = \beta_i^d[u(Y)], \quad i = 1, \ldots, n.$$

Invariance to monotonic transformations is a useful property in decision analysis, and specifically for investment decisions. If $u(y)$ is any von Neumann–Morgenstern (vN-M) utility function, then for all i, $\beta_i^{Ku}(Y) = \beta_i^{Ku}[u(Y)]$ and $\beta_i^{KS}(Y) = \beta_i^{KS}[u(Y)]$. When Y is a NPV, $u(NPV)$ is the only output measure compatible with preferences for the early arrival of cash and the monotonicity of u (Baucells and Sarin 2007). Therefore, the ranking obtained using a risk-neutrality assumption is maintained for any vN-M utility. This allows us to identify the key drivers of uncertainty without having to assess the precise form of $u(y)$.

Another application is in multiattribute utility assessments. Suppose y is a value function over n attributes. Such a function captures the trade-offs between the different attributes under certainty. Risk preferences may be specified by means of a utility function u over values of y, leading to $u(y)$. The choice may involve uncertainty on the attribute values of the alternatives. Sensitivity analysis over such uncertainty can be done without specifying the function u.

In evaluating investment strategies, the final wealth, the return, or the log return can be used as a measure of performance. If Y is the final wealth, Y_0 the initial wealth, $R = (Y - Y_0)/Y$ the return, and $L = \ln(Y/Y_0)$ the log return of the investment, then

$$\beta_i^d(Y) = \beta_i^d(R) = \beta_i^d(L).$$

Invariance to monotonic transformations has computational benefits. Analysts often adopt rescaling to reduce numerical instability when a model output spans several orders of magnitudes (Castaings et al. 2010). The most common form is a logarithmic transformation of the values of y. Invariance to monotonic transformation allows us to perform these transformations directly and still obtain the same β^d. Scale invariance is very useful in engineering applications because analysts often change units or use log transformations of the output (Iman and Hora 1990).

Finally, the economic theory of the consumer is based on ordinal utility. Preference models based on ordinal utility are invariant to monotonic transformations of the utility function (Kreps 1988). The measure β^d preserves this invariance and is appropriate for such models. For example, the Cobb–Douglas utility function over two consumption goods, $y(\theta_1, \theta_2) = \theta_1^{\alpha}\theta_2^{1-\alpha}$, is equivalent to $\ln[y(\theta_1, \theta_2)] = \alpha \ln \theta_1 + (1-\alpha)\ln(\theta_2)$. If we use β^d, the answer of whether θ_1 is more important than θ_2 is the same whether the decision maker uses $y(\cdot)$, $\ln y$, or any other monotonic transformation of $y(\cdot)$.

To compare, variance-based and correlation-based sensitivity measures exhibit invariance to affine transformations only (e.g., changes in units).

4.1. Caveats of Scale Invariance

Scale invariance may not be relevant in some applications. For example, suppose we wish to adjust for the

[5] The calculations can be carried out using symbolic integration software (for the Pareto example, one exploits Equation (11) in Ramsay 2008) or numerically using the estimation algorithm in Appendix B.

[6] Let $U = u(Y)$. Indeed, if $u(y)$ is monotonically increasing, then $F_Y(y) = P(Y \le y) = P(u(Y) \le u(y)) = F_U(u)$, leading to $F_Y(y) - F_{Y|\mathbf{\Theta}_i=\theta_i}(y) = F_U(u) - F_{U|\mathbf{\Theta}_i=\theta_i}(u)$. This also leads to $w[F_Y(y), F_{Y|\mathbf{\Theta}_i=\theta_i}(y)] = w[F_U(u), F_{U|\mathbf{\Theta}_i=\theta_i}(u)]$ so that the d^{MS} family is invariant to monotonically increasing transformations. If $u(y)$ is monotonically decreasing, then $F_U(u) = P(u(Y) \le u(y)) = P(Y \ge y) = 1 - F_Y(y)$, leading to $F_U(u) - F_{U|\mathbf{\Theta}_i=\theta_i}(u) = F_{Y|\mathbf{\Theta}_i=\theta_i}(y) - F_Y(y)$. Because Equations (5) and (6) are symmetric in $F_{Y|\mathbf{\Theta}_i=\theta_i}(y) - F_Y(y)$, then $d^{Ku}\{F_Y, F_{Y|\mathbf{\Theta}_i=\theta_i}\} = d^{Ku}\{F_Y, F_{Y|\mathbf{\Theta}_i=\theta_i}\}$ and $d^{KS}\{F_Y, F_{Y|\mathbf{\Theta}_i=\theta_i}\} = d^{KS}\{F_U, F_{U|\mathbf{\Theta}_i=\theta_i}\}$ also when $u(y)$ is monotonically decreasing. The d^{MS} family is not generally invariant to monotonically decreasing transformations.

effect of the timing of the resolution of uncertainty in investment analyses. One way to do this is using Smith's (1998) framework, which assumes exponential utility for period consumption and derives an *effective NPV* that accounts for the timing of resolution of uncertainty. Effective NPV depends on the utility of consumption and cannot be transformed by a vN-M utility. Of course, β^d is a well-defined way to measure the sensitivity of the effective NPV to different inputs. It just so happens that the property of scale invariance has no use in this context.

Scale invariance is a strong property and assumes that Y has no cardinal interpretation. This may produce some counterintuitive answers. For instance, consider $Y = \mathbf{\Theta}_1 + \phi\mathbf{\Theta}_2$, where $\mathbf{\Theta}_1$ is a 0–1 Bernoulli random variable with $P(\mathbf{\Theta}_1 = 0) = p_0$ and $\mathbf{\Theta}_2$ is uniformly distributed on interval $[0, 1]$. Then it can be shown that

$$\beta_1^{Ku} = 2p_0(1-p_0) \quad \text{and} \quad \beta_2^{Ku} = \max(p_0, 1-p_0). \tag{8}$$

To compare, $\beta_1^{KS} = \beta_1^{Ku}$ (because of dominance) and $\beta_2^{KS} = \frac{3}{4}\beta_2^{Ku}$. Both β^{Ku} and β^{KS} depend on p_0, but they do not depend on ϕ, provided $0 < \phi \leq 1$. To explain this, we note that the support of the unconditional risk profile is $[0, \phi] \cup [1, \phi+1]$ and has no mass on $(\phi, 1)$. By scale invariance, the distance between $u(\phi)$ and $u(1)$ can be modified at will by an appropriate choice of monotonic transformation u. If a sensitivity measure is scale invariant, it is supposed to be insensitive to the amplitude of the gap and hence to the value of ϕ. If $\phi \geq 1$, then the gap disappears, and both β^{Ku} and β^{KS} vary smoothly with ϕ. If one wishes to have a measure that exhibits continuous variations in model parameters at all times, then one could still use metrics based on risk profile separation, such as the L_p-metric class, $d\{F_Y, F_{Y|\mathbf{\Theta}_i=\theta}\} = (\int(|F_Y(y) - F_{Y|\mathbf{\Theta}_i=\theta}(y)|)^p\, dy)^{1/p}$, $1 \leq p < \infty$. These metrics are not scale invariant. In the remainder of this paper, we adhere to scale invariance and adopt the Kuiper metric in our examples and numerical calculations.

5. Uncertain Cash Flows

Throughout this section, we assume the output is the NPV of a discounted cash flow. Unless stated otherwise, uncertain inputs are cash flows. Accordingly,

$$Y = \sum_{l=0}^{L} \frac{\mathbf{\Theta}_l}{(1+r_l)^l} = \mathbf{\Theta} \cdot \phi^T, \tag{9}$$

in which $\phi = [1,\ 1/(1+r_1), \ldots, 1/(1+r_L)^L]$ is the vector of discount factors ($\phi > 0$), r is the cost of capital vector, and $\mathbf{\Theta}$ denotes the random cash flows.

The determination of the NPV's risk profile given the distribution of the cash flows received extensive attention in the 1960s and 1970s (Hillier 1963, Hertz 1964, Barnes et al. 1978).[7] Previous research obtained the unconditional risk profile usually stating an independence assumption. This, in turn, allows us to find the NPV density through the Laplace transform (Grubbström 1967). If the random inputs admit a generic density, $f_{\mathbf{\Theta}}(\mathbf{\Theta})$, then by the change of variable rule, any model of the form $Y = \sum_{i=1}^n a_i \mathbf{\Theta}_i$ yields the following integral expression (Esipenko and Shchuko 1995, Pratt et al. 1995):

$$f_Y(y) = \int \cdots \int f_{\mathbf{\Theta}}(y - a_2 s_1 - a_3 s_2 - \cdots - a_n s_n, s_2, s_3, \ldots, s_n) \prod_{l=1}^{n} \mathrm{d}s_l, \tag{10}$$

and

$$f_{Y|\mathbf{\Theta}_i=\theta_i}(y) = \int \cdots \int f_{\mathbf{\Theta}}(y - \theta_i a_i - a_1 s_1 - a_2 s_2 - \cdots - a_L, s_L, s_1, s_2, \ldots, s_n) \prod_{\substack{l=1 \\ l\neq i}}^{n} \frac{\mathrm{d}s_l}{|a_l|}. \tag{11}$$

The CDFs $F_Y(y)$ and $F_{Y|\mathbf{\Theta}_i=\theta_i}(y)$, required in β_i^{Ku} and β_i^{KS}, are then obtained by integrating Equations (10) and (11), respectively.

5.1. Uniform Two-Period Case

The uniform case is a good starting point to familiarize the reader with the new proposed measure, because it allows us to obtain simple closed-form expressions.

Example 3. Let $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ be two independent uniform random variables on $[0, 1]$, and let $Y = \mathbf{\Theta}_1 + \phi\mathbf{\Theta}_2$, $0 \leq \phi \leq 1$. Then

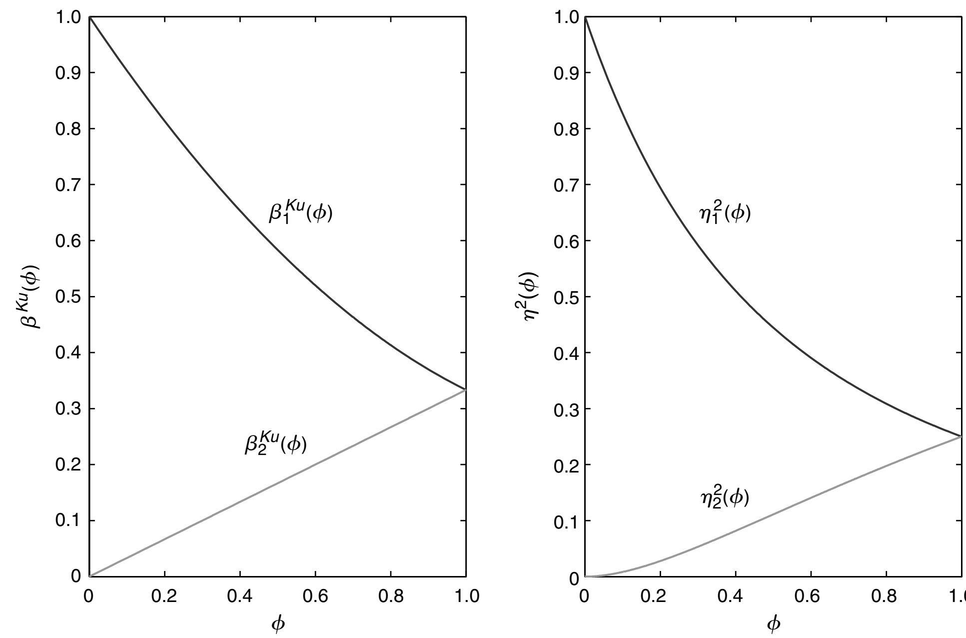
$$\beta_1^{Ku} = \frac{1}{3}\phi^2 - \phi + 1, \quad \beta_2^{Ku} = \frac{1}{3}\phi,$$
$$\eta_1^2 = \frac{1}{1+\phi^2}, \quad \text{and} \quad \eta_2^2 = \frac{\phi^2}{1+\phi^2}.$$

Figure 2 displays the values of these sensitivity measures as a function of ϕ. The importance of $\mathbf{\Theta}_1$ is greater than the importance of $\mathbf{\Theta}_2$. As expected, $\beta_1^{Ku}(\phi)$ decreases with ϕ and $\beta_2^{Ku}(\phi)$ increases with ϕ. By symmetry, they attain the same value at $\phi = 1$; $\eta_1^2(\phi)$ and $\eta_2^2(\phi)$ behave similarly. Observe that $\eta_1^2(\phi) + \eta_2^2(\phi) = 1$ for all ϕ, a property that holds every time Y is additive and the exogenous variables are independent.

This simple example allows us to also draw comparisons with tornado diagrams. Selecting any $\theta^0 = (\theta_1^0, \theta_2^0)$ as a base case, and 0 and 1 as extremes for the variations of $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$, one obtains $TD_1 = y(1, \mathbb{E}[\mathbf{\Theta}_2]) - y(0, \mathbb{E}[\mathbf{\Theta}_2]) = 1$ and $TD_2 = y(\mathbb{E}[\mathbf{\Theta}_1], 1) - y(\mathbb{E}[\mathbf{\Theta}_1], 0) = \phi$, respectively. In contrast with β_1^{Ku} and η_1^2, which decrease with ϕ, TD_1 remains constant.

[7] Recent work addresses the definition of internal rate of return in the presence of stochastic cash flows (Hazen 2009).

Figure 2 $\beta^{Ku}(\phi)$ and $\eta^2(\phi)$ as a Function of the Discount Factor for the Uniform Case



These results show that for all three methods, the ranking of $\mathbf{\Theta}_1$ and $\mathbf{\Theta}_2$ is the same. In this example, the difference between η_2^2 and η_1^2 depends on ϕ^2. In contrast, the difference between β_2^{Ku} and β_1^{Ku} is approximately linear in ϕ, specially for small values of ϕ. The quadratic nature of variance-based measures tends to produce large separations between factor importance. Typically, the most important factor will be much more important than the second, which in turn will be much more important than the third, and so on. This quadratic separation between factors is undesirable because it may lead the decision maker to ignore second- or third-ranked factors that appear irrelevant (say, 1/9 of the most important factor) but actually have a nonnegligible effect (say, 1/3 of the most important factor). Correlation-based and distribution-based measures would not exhibit this behavior.

5.2. Normal Multiperiod Case

Consider a discounted cash flow model (Equation (9)) and let $\mathbf{\Theta} \sim N(\mathbf{m}, \Sigma)$, where $N(\mathbf{m}, \Sigma)$ denotes a multivariate normal distribution with expected values $\mathbf{m}$ and nondegenerate covariance matrix Σ ($\det \Sigma \neq 0$). Let $N(\tau; m, \sigma^2)$ be the normal CDF.

Then[8]

$$\beta_i^{Ku} = \mathbb{E}[\max\{A(\theta_i), B(\theta_i)\} + \max\{C(\theta_i), D(\theta_i)\}], \quad (12)$$

where

$$A(\theta_i) = N(\tau_1(\theta_i); m_Y, \sigma_Y^2) - N(\tau_1(\theta_i); m_{Y|\mathbf{\Theta}_i=\theta_i}, \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i}),$$
$$B(\theta_i) = N(\tau_1(\theta_i); m_{Y|\mathbf{\Theta}_i=\theta_i}, \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i}) - N(\tau_1(\theta_i); m_Y, \sigma_Y^2),$$
$$C(\theta_i) = N(\tau_2(\theta_i); m_Y, \sigma_Y^2) - N(\tau_2(\theta_i); m_{Y|\mathbf{\Theta}_i=\theta_i}, \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i}),$$
$$D(\theta_i) = N(\tau_2(\theta_i); m_{Y|\mathbf{\Theta}_i=\theta_i}, \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i}) - N(\tau_2(\theta_i); m_Y, \sigma_Y^2),$$

$$\tau_{1,2}(\theta_i) = \Bigg(\sigma_Y^2 m_{Y|\mathbf{\Theta}_i=\theta_i} - \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i} m_Y \pm \sqrt{\sigma_Y^2 \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i}\left[(\phi_i\theta_i)^2 + (\sigma_Y^2 - \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i})\ln\left(\frac{\sigma_Y^2}{\sigma^2_{Y|\mathbf{\Theta}_i=\theta_i}}\right)\right]}\Bigg) \cdot (\sigma_Y^2 - \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i})^{-1},$$

$$\sigma_Y^2 = \boldsymbol{\phi}\Sigma\boldsymbol{\phi}^T, \quad \sigma^2_{Y|\mathbf{\Theta}_i=\theta_i} = \boldsymbol{\phi}\Sigma_{Y|\mathbf{\Theta}_i=\theta_i}\boldsymbol{\phi}^T, \quad m_Y = \boldsymbol{\phi}\mathbf{m}$$

$$m_{Y|\mathbf{\Theta}_i=\theta_i} = \sum_{s=0}^{L} \phi_s \left[\mathbb{E}[\theta_s] + (\theta_i - \mathbb{E}[\theta_i])\frac{\rho_{s,i}\sigma_s}{\sigma_i}\right], \quad i = 1, 2, \ldots, n, \quad \text{and}$$

$$\Sigma_{Y|\mathbf{\Theta}_i=\theta_i} = \left[\sigma_{j,s} - \frac{\sigma_{j,i}\cdot\sigma_{i,s}}{\sigma_i}, \; j, s, i = 1, 2, \ldots, n\right].$$

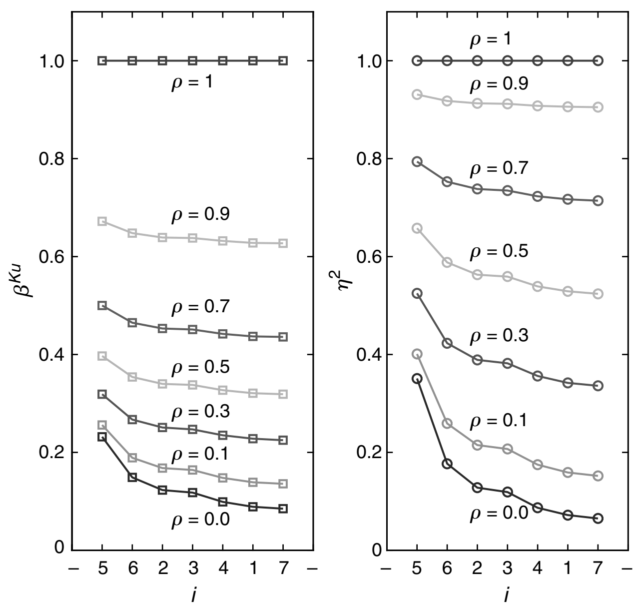
Also,

$$\eta_i^2 = \frac{\mathbb{V}[\mathbf{\Theta}_i] + 2\sum_{k \neq i=0}^{L} \mathrm{Cov}[\mathbf{\Theta}_k, \mathbf{\Theta}_i]/[(1+r_i)^i(1+r_k)^k]}{\mathbb{V}[Y]}. \quad (13)$$

As an example, consider the NPV model in Beccacece et al. (2001), developed to study an investment in the energy sector. The requisite discounted

[8] Because the unconditional and conditional distributions are normal, and the normal distribution is continuous and differentiable, the $\sup_y\{F_Y(y) - F_{Y|\mathbf{\Theta}=\theta_i}(y)\}$ is attained at the points where $(d/dy)(F_Y(y) - F_{Y|\mathbf{\Theta}=\theta_i}(y)) = 0$, which we denote by $\tau_{1,2}(\theta_i)$ (Borgonovo et al. 2011, Prop. 4). Substituting into the expression of the Kuiper metric completes the calculations.

Figure 3 Pareto Chart with β_i^{Ku} (Left) and η_i^2 (Right) for the Seven Cash Flows of Example 4 with Correlation Levels $\rho = 0, 0.1, 0.3, 0.5, 0.7, 0.9$, and 1



cash flow model considers seven periods; Periods 1–4 are construction periods; periods 5–7 are operations periods. Cash flows are normally distributed, and we consider possible correlations among them.

Example 4. Let $\boldsymbol{\Theta}_i$, $i = 1, \ldots, 7$, be normally distributed. The means are $\boldsymbol{\mu} = (-10, -30, -50, -70, 600, 750, 900)$, the variances are $\boldsymbol{\sigma}^2 = (3.2, 5.5, 7.1, 8.4, 24.5, 27.4, 30)$, and the discount factors are $\boldsymbol{\phi} = (0.95, 0.73, 0.55, 0.40, 0.27, 0.17, 0.096)$. Let $\rho \in [0, 1]$ be the correlation between all pairs.

We compute the sensitivity measures for $\rho \in \{0, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$. We obtain β_i^{Ku} and η_i^2 by implementing Equations (13) and (12) using symbolic mathematical software, without resorting to MCS. Figure 3 reports the results. For both β_i^{Ku} and η_i^2, we find that $\boldsymbol{\Theta}_5$ is the most important cash flow, followed by $\boldsymbol{\Theta}_6$, $\boldsymbol{\Theta}_2$, $\boldsymbol{\Theta}_3$, $\boldsymbol{\Theta}_4$, $\boldsymbol{\Theta}_1$, and $\boldsymbol{\Theta}_7$.

If $\rho = 0$, then $\sum_{i=1}^{7} \eta_i^2 = 1$ and η_i^2 represent the fraction of $\mathbb{V}[Y]$ contributed by $\boldsymbol{\Theta}_i$. This interpretation is lost when $\rho > 0$ because of dependencies (Oakley and O'Hagan 2004). In contrast, the interpretation of β^{ku} is the same for any correlation level.

Both β_i^{Ku} and η_i^2 systematically increase as the correlation increases (see Figure 3). The result is intuitive. In the presence of dependencies, the fact that $\boldsymbol{\Theta}_i = \theta_i$ resolves uncertainty in $\boldsymbol{\Theta}_i$ and provides information on the cash flows that are correlated with that input. At $\rho = 1$ all inputs are equally important because of perfect correlation.

5.3. Idiosyncratic Risks

The model in Beccacece et al. (2001) is built for taking idiosyncratic risks into account. In particular, one can consider the plant rejection risk at the last year of construction.

Example 4 (Continued). Let $\boldsymbol{\Theta}_8$ denote the random variable used to account for plant rejection risk. The cash flow stream then can be written as

$$Y = \sum_{i=1}^{4} \phi_i \boldsymbol{\Theta}_i + \boldsymbol{\Theta}_8 \left[\sum_{i=5}^{7} \phi_i \boldsymbol{\Theta}_i \right], \tag{14}$$

with

$$\boldsymbol{\Theta}_8 = \begin{cases} 1 & \text{with } p = 0.6, \\ -1.5 & \text{with } p = 0.4, \end{cases}$$

where the -1.5 accounts for the fact that the plant sponsor will suffer further losses related to dismantling costs, along with the missed cash flows. The remaining cash flow distributions are as before.

We can also obtain variance-based sensitivity measures in this case (see Appendix A for the calculations). We report their values in the second row of Table 2.

Observe that in this example variance-based sensitivity measures provide no insight. The cash flows are judged to have very low relevance. $\boldsymbol{\Theta}_5$, $\boldsymbol{\Theta}_6$, and $\boldsymbol{\Theta}_7$, the key uncertainty drivers in the previous version of the model, would be considered uninfluential because their variance-based importance is exactly equal to 0. A decision maker would think that the only uncertain factor on which to focus is $\boldsymbol{\Theta}_8$. Instead, β^{Ku} shows that $\boldsymbol{\Theta}_8$ is the most important factor, but it also allows us to appreciate that $\boldsymbol{\Theta}_5$ still has a strong effect, and it is followed still by $\boldsymbol{\Theta}_6$ and $\boldsymbol{\Theta}_7$, with the ranking of the remaining random inputs being unaltered with respect to the original version of the problem. Thus, for this example, β^{Ku} provides a more reasonable indication than variance-based sensitivity measures.

5.4. Case Study: The Genzyme/GelTex Joint Venture

In general applications, the inputs may not be normally distributed and may contain a mixture of discrete and continuous variables. β^{Ku} is not analytically known and needs to be estimated using MCS. We illustrate this with a real-world case study, "Genzyme/GelTex Pharmaceuticals Joint Venture" (Jacquet et al. 1999).

GelTex is a biotech company created to market the innovative research ideas of the Whitesides Research Group of Harvard University. One of the new GelTex products, RenaGel, has entered the final approval stages of the U.S. Food and Drug Administration (FDA), having passed the first four steps needed for final FDA approval. As is well known, the commercialization of a new drug follows FDA protocol. Given the low probability of a new drug passing the first

Table 2 Exogenous Variables' Importance with Explicit Modelling of Plant Rejection Risk for Example 4

Variable	Θ_1	Θ_2	Θ_3	Θ_4	Θ_5	Θ_6	Θ_7	Θ_8
η^2	1.341E–5	1.372E–5	9.916E–6	6.107E–6	0	0	0	0.999956
β^{Ku}	0.0579	0.0775	0.0747	0.0636	0.2436	0.1594	0.0952	0.4

four of the six steps needed to get final approval, RenaGel is an attractive investment proposition.

RenaGel has agreements with other partners to commercialize the product in Asia. Genzyme is the fourth largest biotech company in the United States. Genzyme is investigating whether to enter a joint venture with GelTex for the U.S. and European sales of RenaGel. Genzyme needs to determine the value of investing in RenaGel.

A financial model calculating the NPV of RenaGel is provided as part of the case. The first column of Table 3 contains the list of relevant uncertainties. The first is whether RenaGel will be finally approved by the FDA. If approval fails, then Genzyme undergoes a net loss of \$16 million. The probability of approval, judging by historical data, is estimated at 65%.

The remaining uncertainties are conditional on approval. The second concerns the timing of the drug's launch. Analysts consider three scenarios: no delay, a one-year delay, and a two-year delay. This is a discrete random variable, and probabilities are provided. The other factors are peak penetration rate, price per patient, compliance rate, gross profit, marketing cost multiplier, and life of the drug. These factors are continuous random variables. For each, three values are provided: low, base, and high. The case suggests the use of triangular distributions with low and high end points and the base as the most likely value, denoted by $\Delta(low; most\_likely; high)$.

The case provides a spreadsheet that produces the NPV. The NPV depends on the eight exogenous variables displayed in Table 3 with the corresponding distributions.

The business plan allows one to investigate the profitability of the deal. Setting the base case value of each exogenous variable equal to its most likely value (let us denote this point as $\boldsymbol{\theta}^0$), we obtain $Y(\boldsymbol{\theta}^0) \simeq$ 90 million. Here, the FDA approval variable is set to approval. The base scenario set to nonapproval yields a certain NPV loss of \$16 million.

Table 3 Uncertain Inputs in the Genzyme/GelTex Business Case

Variable	Probability distribution	β^{Ku}	η^2
Θ_1 *FDA approval*	Boolean with $p_{approval} = 0.65$	0.4550	0.5719
Θ_2 *Launch delay* (years)	$0y, 0.7;\ 1y, 0.2;\ 2y, 0.1$	0.0668	0.0257
Θ_3 *Peak penetration rate* (%)	$\Delta(0.2; 0.5; 0.59)$	0.1791	0.1086
Θ_4 *Price per patient* (\$)	$\Delta(600; 1{,}100; 1{,}300)$	0.1235	0.0605
Θ_5 *Compliance* (%)	$\Delta(0.75; 0.92; 0.94)$	0.0446	0.0067
Θ_6 *Gross profit* (%)	$N(0.7; 0.05)$	0.0705	0.0214
Θ_7 *Marketing cost multiplier*	$\Delta(0.87; 0.93; 1.2)$	0.0248	0.0002
Θ_8 *Life of drug*	Poisson with mean 14	0.0982	0.0448

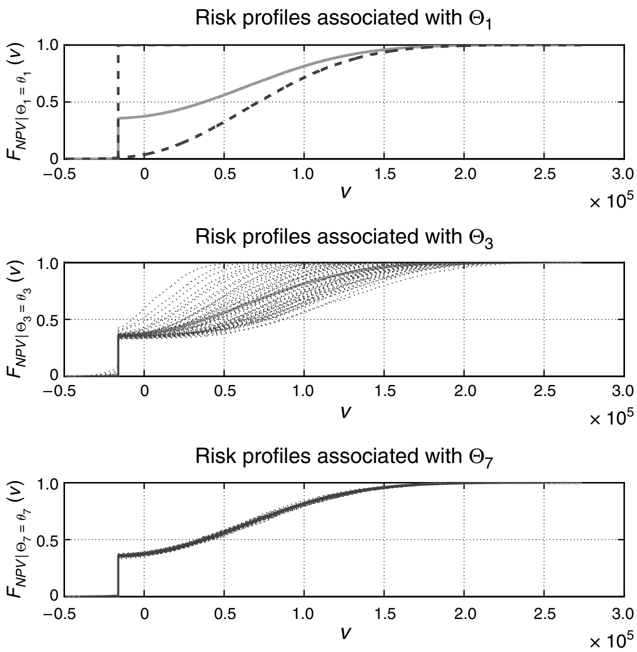
A key learning point of the case is the need to calculate expected NPV explicitly, rather than burying uncertainty in the discount rate and calculating a single NPV. A decision tree approach with two branches readily accounts for the risk of FDA approval. To account for the remaining assumptions, the decision tree approach becomes impractical. The natural way to proceed is to employ MCS to generate random scenarios and estimate $\mathbb{E}[NPV]$.

To perform the MCS, we utilize Crystal Ball, generating a Latin hypercube sample of 100,000. The analysis takes less than one minute on a personal PC. We let $\hat{Y}$ denote the column vectors containing the NPVs' realizations. Let M denote the computational cost of estimating a global sensitivity statistic. M is measured by number of model evaluations. A brute force estimation of β^{Ku} and η^2 is associated with $M = n \cdot N^2$, where n is the number of exogenous variables (random inputs) and N the sample size of the probabilistic sensitivity analysis. The N^2 term is generated by the need to perform N additional Monte Carlo simulations at each of the N sampled values of $\boldsymbol{\Theta}_i$. Plischke (2010) and Plischke et al. (2013), however, prove that both η^2 and any distribution-based sensitivity measure can be obtained at the cost $M = N$ by postprocessing the results of a usual probabilistic sensitivity, with no additional model evaluations. We use this same approach to obtain β^{Ku} directly from the data set $\hat{\Psi}_{N\times n+1} = [\hat{\boldsymbol{\Theta}}\hat{Y}]$, where $\hat{\boldsymbol{\Theta}}$ is the $N \times n$ model input sample. We describe the algorithm in Appendix B and apply it to $\hat{\Psi}_{100{,}000\times 9}$ and obtain the results in Figure 4. The same approach could be used to calculate β^{KS}.

The unconditional and conditional risk profiles displayed in Figure 4 allow us to appreciate the impact of a variable on the joint venture risk profile. Let us consider the top graph, displaying the impact of $\boldsymbol{\Theta}_1$, *FDA approval*. The continuous line represents the unconditional risk profile. One notes that NPV is a mixed random variable, with an atom at $Y = -16M$. The dotted lines represent the conditional risk profiles given $\boldsymbol{\Theta}_1 = 1$ (FDA approval) and $\boldsymbol{\Theta}_1 = 0$ (no FDA approval). $\boldsymbol{\Theta}_1 = 0$ makes the risk profile a step function centered at $-16M$. In fact, because $\boldsymbol{\Theta}_1$ is a discrete Bernoulli variable, we can write

$$\beta_1^{Ku} = P(\boldsymbol{\Theta}_1 = 0) d^{Ku}_{\boldsymbol{\Theta}_1=0} + P(\boldsymbol{\Theta}_1 = 1) d^{Ku}_{\boldsymbol{\Theta}_1=1}$$
$$= 0.35 \times 0.65 + 0.65 \times 0.35 = 0.455.$$

Figure 4 **Unconditional and Conditional Risk Profiles Associated with Θ_1, Θ_3, and Θ_7**



To see this, note that (i) given $\mathbf{\Theta}_1 = 0$, we have $\Delta P^s = 0$ and $\Delta P^f = 0.65$, which is the maximum separation between the conditional and unconditional distributions occurring at $-16M$, thus $d^{Ku}_{\mathbf{\Theta}_1=0} = 0.65$; and (ii) given $\mathbf{\Theta}_1 = 1$, the maximum separation still occurs at $-16M$, with $d^{Ku}_{\mathbf{\Theta}_1=1} = 0.35$.

The middle graph in Figure 4 displays the impact of $\mathbf{\Theta}_3$. $\mathbf{\Theta}_3$ is a continuous random variable and leads to the continuous set of risk profiles displayed in Figure 4. The bottom graph displays the conditional risk profiles obtained when fixing $\mathbf{\Theta}_7$. Little or no variation in the risk profile is registered. We therefore expect a small value of β^{Ku}_7.

The values of β^{Ku} are visualized in the second-to-last column of Table 3. The most influential factor is $\mathbf{\Theta}_1$ (*FDA approval*), followed by $\mathbf{\Theta}_3$ (*peak penetration rate*), $\mathbf{\Theta}_4$ (*price per patient*), and $\mathbf{\Theta}_8$ (*life of the drug*). The least influential factors are $\mathbf{\Theta}_5$ (*marketing cost multiplier*) and $\mathbf{\Theta}_7$ (*compliance rate*). The remaining exogenous variables have an intermediate influence.

The results of Table 3 suggest that if resources are available for further modelling or data collection, the decision maker should prioritize that on FDA approval. Moreover, conditional on approval, managerial efforts should focus on reaching the penetration market rate and on extending as much as possible the life of the drug. Also, an analyst equipped with these results probably ought not to invest a considerable amount of time trying to accurately model the marketing cost multiplier or the gross profit percentage at this stage of the investigation.

Using the algorithm of Plischke (2010), we estimate the variance-based sensitivity measures from $\hat{\Psi}$. The results are reported in the last column of Table 3. It is reassuring that the ranking of the factors of both β^{Ku} and η^2 coincide. An analyst can use η^2 to extract additional insights. Because we assume independence of the random variables, the sum $\sum_{i=1}^{8} \eta_i^2$ coincides with the fraction of the model variance that can be explained by individual effects. In our case, $\sum_{i=1}^{8} \hat{\eta}_i^2 \simeq 0.84$. This indicates that the model responds almost additively to the exogenous variable variations ($\mathbf{\Theta}_1$ has a strong individual effect), but interaction effects cannot be neglected.

Using η^2, note that the relative importance of the inputs decays very rapidly. According to η^2, peak penetration rate, price per patient, and life of the drug are much less important than is FDA approval. In contrast, β^{Ku} gives a more balanced picture, one where these variables have a sizeable importance relative to FDA approval. Again, we believe this is an advantage of β^{Ku} over η^2.

We have carried out the sensitivity analysis in a risk-neutral setting. If one were to introduce the decision maker's utility function, u, and compute $u(Y)$, then the values of η^2 would need to be recalculated, whereas the values of β^{Ku} would remain unchanged.

6. Conclusion

We have proposed a new class of sensitivity measures based on scale-invariant distances between risk profiles. From a decision analysis perspective, scale invariance makes the results independent of the utility function. In addition, the sensitivity measures are responsive to nonlinearities in the model, they are well defined in the presence of probabilistic dependencies among inputs, they do not require additional computations with respect to a standard Monte Carlo simulation, and they do not require additional information from that which is standard in investment analysis with uncertainties.

Among all scale-invariant sensitivity measures, we propose the one based on the Kuiper metric and consider as well the companion measure based on the Kolmogorov–Smirnov metric. Both metrics admit an interpretation of probabilistic influence of each input on the output and bypass limitations of contribution to variance and other existing measures. We have provided closed-form expressions for β^{Ku} in particular cases. β^{Ku} (or β^{KS}) can be estimated numerically as a by-product of an MCS. If implemented in MCS software, it would provide users with a sound method for identifying key uncertainty drivers.

Acknowledgments

The authors thank the associate editor (anonymous), the department editor (Peter Wakker), and the anonymous referees for the very perceptive comments. Financial support from the ELEUSI Research Centre of Bocconi University is gratefully acknowledged by Emanuele Borgonovo. Financial support from the Ministerio de Ciencia e Innovación [Grant ECO2010-17145] is gratefully acknowledged by Manel Baucells.

Appendix A. Proofs

Proof of Lemma 1. By $F_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = \prod_{i=0}^n F_{\boldsymbol{\Theta}_i}(\theta_i)$,

$$\mathbb{E}[Y \mid \boldsymbol{\Theta}_i = \theta_i] = g_i(\theta_i) \prod_{l=1,\, l\neq i}^{n} \mathbb{E}[g_l(\boldsymbol{\Theta}_l)] + K \quad \forall\, i. \tag{A1}$$

Because $\mathbb{E}[g_l(\boldsymbol{\Theta}_l)] = 0$, we have $\mathbb{E}[Y \mid \boldsymbol{\Theta}_i = \theta_i] = K$. Consequently, $\mathbb{V}\{\mathbb{E}[Y \mid \boldsymbol{\Theta}_i = \theta_i]\} = 0$. Thus, the numerator in Equation (2) is null $\forall\, i$, which proves that $\eta_i^2 = 0\, \forall\, i$. By independence and by the fact that $\mathbb{E}[g_i(\boldsymbol{\Theta}_i)] = 0\ \forall\, i$, the variance of Y is given by

$$\mathbb{V}[Y] = \prod_{i=1}^{n} \mathbb{E}[g_i(\theta_i)^2]. \tag{A2}$$

Now, consider the numerator in the second term on the right-hand side of Equation (3), $\mathbb{V}\{\mathbb{E}[Y \mid \boldsymbol{\Theta}_{\sim i}]\}$. We have $\mathbb{E}[Y \mid \boldsymbol{\Theta}_{\sim i}] = \prod_{l=1,\, l\neq i}^{n} g_l(\theta_l)\mathbb{E}[g_i(\boldsymbol{\Theta}_i)] = 0\ \ \forall\, \boldsymbol{\Theta}_{\sim i}$. Thus, $\mathbb{V}\{\mathbb{E}[Y \mid \boldsymbol{\Theta}_{\sim i}]\} = 0$, and by Equation (3), $\eta_i^T = 1\ \forall\, i$. Q.E.D.

Proof of Proposition 1. If

$$d\{F_{u(Y)}(u), G_{u(Y)}(u)\} = d\{F_Y(y), G_Y(y)\}, \tag{A3}$$

then for all $\boldsymbol{\Theta}_i = \theta_i$,

$$d\{F_Y, F_{Y\mid \boldsymbol{\Theta}_i=\theta_i}\} = d\{F_{u(Y)}, F_{u(Y)\mid \boldsymbol{\Theta}_i=\theta_i}\}. \tag{A4}$$

Hence, when taking the expectation of both sides with respect to the distribution of $\boldsymbol{\Theta}_i$, one obtains

$$\mathbb{E}[d\{F_Y, F_{Y\mid \boldsymbol{\Theta}_i=\theta_i}\}] = \mathbb{E}[d\{F_{u(Y)}, F_{u(Y)\mid \boldsymbol{\Theta}_i=\theta_i}\}]. \tag{A5}$$

This last equality is equivalent to saying that $\beta_i^d(Y) = \beta_i^d[u(Y)]$. Q.E.D.

Proof of Example 3. We have

$$F_Y(v) = \begin{cases} \dfrac{v^2}{2\phi}, & 0 \le v < \phi; \\[2ex] v - \dfrac{\phi}{2}, & \phi \le v < 1; \\[2ex] \dfrac{1}{2\phi}(-\phi^2 + 2\phi v - v^2 + 2v - 1), & 1 \le v < 1+\phi. \end{cases} \tag{A6}$$

Suppose that the decision maker is informed that $\boldsymbol{\Theta}_1 = \theta_1$. Then

$$F_{Y\mid \boldsymbol{\Theta}_1}(v \mid \theta_1) = \begin{cases} 0, & v < \theta_1; \\[2ex] \dfrac{v}{\phi} - \dfrac{\theta_1}{\phi}, & \theta_1 < v < \theta_1 + \phi; \\[2ex] 1, & v > \theta_1 + \phi. \end{cases} \tag{A7}$$

Similarly,

$$F_{Y\mid \boldsymbol{\Theta}_2}(v \mid \theta_2) = \begin{cases} 0, & v < \phi\theta_2; \\ v - \phi\theta_2, & \phi\theta_2 < v < \phi\theta_2 + 1; \\ 1, & v > \phi\theta_2 + 1. \end{cases} \tag{A8}$$

From this, we find the following separations:

$$\begin{aligned} d_1^{Ku}(\theta_1) &= F_Y(\theta_1) + 1 - F_Y(\theta_1 + \phi), \\ d_2^{Ku}(\theta_2) &= F_Y(\theta_2 \cdot \phi) + 1 - F_Y(\theta_2 \cdot \phi). \end{aligned} \tag{A9}$$

Consider now $d_1^{Ku}(\theta_1, \phi)$. By Equation (A6), if $\phi \le \frac{1}{2}$, then

$$d_1^{Ku}(\theta_1, \phi) = \begin{cases} \dfrac{1}{2\phi}(-\phi^2 - 2\phi\theta_1 + 2\phi + \theta_1^2) & \text{if } \theta_1 < \phi, \\[2ex] 1 - \phi & \text{if } \phi < \theta_1 < 1 - \phi, \\[2ex] \dfrac{1}{2\phi}(-\phi^2 + 2\phi\theta_1 + \theta_1^2 - 2\theta_1 + 1) & \text{if } 1 - \phi < \theta_1 < 1; \end{cases}$$

and if $\phi > \frac{1}{2}$, then

$$d_1^{Ku}(\theta_1, \phi) = \begin{cases} \dfrac{1}{2\phi}(-\phi^2 - 2\phi\theta_1 + 2\phi + \theta_1^2) & \text{if } \theta_1 < \phi, \\[2ex] \dfrac{1}{2\phi}(2\theta_1^2 - 2\theta_1 + 1) & \text{if } \phi < \theta_1 < 1 - \phi, \\[2ex] \dfrac{1}{2\phi}(-\phi^2 + 2\phi\theta_1 + \theta_1^2 - 2\theta_1 + 1) & \text{if } 1 - \phi < \theta_1 < 1. \end{cases}$$

Then by integration,

$$\beta_1^{Ku} = \int_0^1 d_1^{Ku}(\theta_1, \phi)\, d\theta_1 = \tfrac{1}{3}\phi^2 - \phi + 1$$

for all values of ϕ. For $d^{Ku}(\theta_2)$, we have the following. By Equation (A6),

$$d^{Ku}(\theta_2) = \frac{\phi^2\theta_2^2}{2\phi} + 1 - \left(\frac{1}{2\phi}(-\phi^2 + 2\phi(\phi\theta_2 + 1) - (\phi\theta_2 + 1)^2 + 2(\phi\theta_2 + 1) - 1)\right) = \frac{1}{2}\phi(2\theta_2^2 - 2\theta_2 + 1).$$

Then by integration,

$$\int_0^1 \left(\tfrac{1}{2}\phi(2s^2 - 2s + 1)\right) ds = \tfrac{1}{3}\phi. \quad \text{Q.E.D.}$$

Calculations for Table 2. By Equation (14), and the properties of variance, we have

$$\mathbb{V}[Y] = \sum_{i=1}^{4} \phi_i^2 \mathbb{V}[\boldsymbol{\Theta}_i] + \mathbb{V}\left[\boldsymbol{\Theta}_8\left(\sum_{i=5}^{7} \phi_i \boldsymbol{\Theta}_i\right)\right]. \tag{A10}$$

Let us call $\boldsymbol{\Theta}_8 = C$ and $[\sum_{i=5}^{7} \phi_i \boldsymbol{\Theta}_i] = B$. Then we have

$$\mathbb{V}[CB] = \mathbb{E}[C^2B^2] - \mathbb{E}[CB]^2. \tag{A11}$$

By independence,

$$\mathbb{V}[CB] = \mathbb{E}[C^2]E[B^2] - \mathbb{E}[C]^2\,\mathbb{E}[B]^2. \tag{A12}$$

Because $\mathbb{E}[C] = 0$,

$$\mathbb{V}[CB] = \mathbb{E}[C^2]\,\mathbb{E}[B^2]. \tag{A13}$$

Hence,

$$\mathbb{V}[Y] = \sum_{i=1}^{4} \phi_i^2 \mathbb{V}[\boldsymbol{\Theta}_i] + \mathbb{E}[\boldsymbol{\Theta}_8^2]\,\mathbb{E}\left[\left(\sum_{i=5}^{7} \phi_i \boldsymbol{\Theta}_i\right)^2\right]. \tag{A14}$$

Then for $i=1,\ldots,4$,

$$\eta_i^2=\frac{\phi_i^2\mathbb{V}[\boldsymbol{\Theta}_i]}{\mathbb{V}[Y]}. \tag{A15}$$

For $i=5$, we have

$$[Y\mid\boldsymbol{\Theta}_5=\theta_5]=\sum_{i=1}^{4}\phi_i\boldsymbol{\Theta}_i+\boldsymbol{\Theta}_8\left(\phi_5\theta_5+\sum_{i=6}^{7}\phi_i\boldsymbol{\Theta}_i\right) \tag{A16}$$

and

$$\mathbb{E}[Y\mid\boldsymbol{\Theta}_5=\theta_5]=\sum_{i=1}^{4}\phi_i\mathbb{V}[\boldsymbol{\Theta}_i]+\mathbb{E}[\boldsymbol{\Theta}_8]\left(\phi_5\theta_5+\sum_{i=6}^{7}\phi_i\mathbb{E}[\boldsymbol{\Theta}_i]\right). \tag{A17}$$

Because $\mathbb{E}[\boldsymbol{\Theta}_8]=0$,

$$\mathbb{E}[Y\mid\boldsymbol{\Theta}_5=\theta_5]=\sum_{i=1}^{4}\phi_i\mathbb{V}[\boldsymbol{\Theta}_i]. \tag{A18}$$

Finally, because $\sum_{i=1}^{4}\phi_i\mathbb{V}[\boldsymbol{\Theta}_i]$ is independent of $\boldsymbol{\Theta}_5$,

$$\mathbb{V}(\mathbb{E}[Y\mid\boldsymbol{\Theta}_5=\theta_5])=0. \tag{A19}$$

The same argument applies to $i=6,7$, and we conclude that $\eta_5^2=\eta_6^2=\eta_7^2=0$.

For $i=8$, we have that

$$[Y\mid\boldsymbol{\Theta}_8=\theta_8]=\sum_{i=1}^{4}\phi_i\boldsymbol{\Theta}_i+\theta_8\left(\sum_{i=5}^{7}\phi_i\boldsymbol{\Theta}_i\right), \tag{A20}$$

$$\mathbb{E}[Y\mid\boldsymbol{\Theta}_8=\theta_8]=\sum_{i=1}^{4}\phi_i\,\mathbb{E}[\boldsymbol{\Theta}_i]+\theta_8\left(\sum_{i=5}^{7}\phi_i\,\mathbb{E}[\boldsymbol{\Theta}_i]\right), \tag{A21}$$

and

$$\begin{aligned}\mathbb{V}(\mathbb{E}[Y\mid\boldsymbol{\Theta}_8=\theta_8])&=\mathbb{V}[\theta_8]\,\mathbb{E}\left[\left(\sum_{i=5}^{7}\phi_i\boldsymbol{\Theta}_i\right)\right]^2\\&=\mathbb{E}[\boldsymbol{\Theta}_8^2]\,\mathbb{E}\left[\left(\sum_{i=5}^{7}\phi_i\boldsymbol{\Theta}_i\right)^2\right].\end{aligned} \tag{A22}$$

Hence,

$$\eta_8^2=\frac{\mathbb{E}[\boldsymbol{\Theta}_8^2]\,\mathbb{E}\left[\left(\sum_{i=5}^{7}\phi_i\boldsymbol{\Theta}_i\right)^2\right]}{\mathbb{V}[Y]}. \tag{A23}$$

Substituting the numerical values, one finds the results presented in Table 2. The values of β^{Ku} are obtained numerically. Q.E.D.

Appendix B. Principles of the Estimation Algorithm

Central to the estimation in Plischke et al. (2013) is the notion of class density. Because in this work we are measuring the distance between CDFs, we replace class density with class-empirical CDF. We obtain a class-empirical CDF as follows. Let $[\hat{\boldsymbol{\Theta}}\hat{Y}]_{N\times(n+1)}$ be the input–output data set of the MCS from which the risk profile is estimated. For $i=1,2,\ldots,n$, form the scatterplot $\hat{\boldsymbol{\Theta}}_i-\hat{Y}$. Then partition $\hat{\boldsymbol{\Theta}}_i$ in m classes, $\mathcal{P}=\{\mathcal{C}_m\mid m=1,\ldots,M\}$, $\bigcup_{m=1}^{M}\mathcal{C}_m=\hat{\boldsymbol{\Theta}}_i$, $\mathcal{C}_m\cap\mathcal{C}_{m'}=\varnothing$, $m\neq m'$. Finally, we define the class CDF as the empirical CDF conditional on $\boldsymbol{\Theta}_i\in\mathcal{C}_m$, denoted by $\hat{F}_{Y\mid\boldsymbol{\Theta}_i\in\mathcal{C}_m}(y)$. Equations (5) and (6) are estimated via the statistics

$$\hat{d}_m^{Ku}=\sup_y\{\hat{F}(y)-\hat{F}_{Y\mid\theta_i\in\mathcal{C}_m}(y)\}+\sup_y\{\hat{F}_{Y\mid\theta_i\in\mathcal{C}_m}(y)-\hat{F}(y)\} \tag{B1}$$

and

$$\hat{d}_m^{KS}=d\{F_Y,F_{Y\mid\boldsymbol{\Theta}_i=\theta}\}=\sup_y|\hat{F}(y)-\hat{F}_{Y\mid\theta_i\in\mathcal{C}_m}(y)|, \tag{B2}$$

where $\hat{F}(y)$ is the risk profile. The β_i^{Ku} estimation statistics are, then,

$$\hat{\beta}_i^{Ku}=\sum_{m=1}^{M}\hat{d}_m^{Ku}\cdot\frac{n_m}{n}\quad\text{and}\quad\hat{\beta}_i^{KS}=\sum_{m=1}^{M}\hat{d}_m^{KS}\cdot\frac{n_m}{n}, \tag{B3}$$

where n_m is the number of points in $\mathcal{C}_m$ and n_m/n is a Monte Carlo estimate of the probability that $\boldsymbol{\Theta}_i\in\mathcal{C}_m$, denoted here by $\hat{P}_{\boldsymbol{\Theta}_i}(\mathcal{C}_m)$. Theorem 2 in Plischke et al. (2013) ensures that the estimators in Equation (B3) are consistent.

If $\boldsymbol{\Theta}_i$ is a discrete random variable, then we can exploit our knowledge of the probability distribution of $\boldsymbol{\Theta}_i$ by setting $\hat{P}_{\boldsymbol{\Theta}_i}(\mathcal{C}_m)=p_m$. If $\boldsymbol{\Theta}_i$ is a continuous random variable, then using equipopulated partitions leads to the statistics $\hat{\beta}_i^{Ku}=\hat{d}_m^{Ku}/M$.

Note that our choice of the metric allows one to benefit from the availability of subroutines contained in standard software packages (from Matlab, to R, to @Risk) for estimating the statistic in Equation (B1). In the present work, we create a postprocessing algorithm in Matlab and estimate $\hat{d}_m$ by the kstest2.m function.

For samples of limited size, we obtain confidence intervals in the estimates through the bootstrap method by the distribution of

$$\beta_i^*=2\hat{\beta}_i^{Ku}-\beta_i^{BS}\ (\text{or }\beta_i^*=2\hat{\beta}_i^{KS}-\beta_i^{BS}), \tag{B4}$$

where β_i^{BS} is the estimate of β_i^{Ku} (β_i^{KS}) produced by the nth bootstrap replicate (Efron and Gong 1983).

References

Anderson TW, Darling DA (1952) Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* 23(2):193–212.

Barnes JW, Zinn CD, Eldred BS (1978) A methodology for obtaining the probability density function of the present worth of probabilistic cash flow profiles. *IIE Trans.* 10(3):226–236.

Baucells M, Sarin RK (2007) Evaluating time streams of income: Discounting what? *Theory Decision* 63(2):95–120.

Beccacece F, Gallo P, Peccati L (2001) Survival risk and project evaluation. Ruan D, Kacprzyk J, Fedrizzi M, eds. *Soft Computing for Risk Evaluation and Management* (Physica-Verlag, Heidelberg, Germany), 397–409.

Bedford T (1998) Sensitivity indices for (tree)-dependent variables. *Proc. Second Internat. Sympos. Sensitivity Anal. Model Output* (Office for the Official Publications of the European Commission, Luxembourg, UK), 17–20.

Benninga S (2008) *Financial Modelling With Excel* (MIT Press, Cambridge, MA).

Berkowitz J (2001) Testing density forecasts, with applications to risk management. *J. Bus. Econom. Statist.* 19(4):465–474.

Bordley R, Li Calzi M (2000) Decision analysis using targets instead of utility functions. *Decisions Econom. Finance* 23(1):53–74.

Borgonovo E (2007) A new uncertainty importance measure. *Reliability Engrg. System Safety* 92(6):771–784.

Borgonovo E, Castaings W, Tarantola S (2011) Moment independent uncertainty importance: New results and analytical test cases. *Risk Anal.* 31(3):404–428.

Campolongo F, Saltelli A (1997) Sensitivity analysis of an environmental model: An application of different analysis methods. *Reliability Engrg. System Safety* 57(1):49–69.

Castaings W, Borgonovo E, Tarantola S (2010) Sampling plans for the estimation of moment-independent importance measures. *SAMO 2010: Proc. Sixth Internat. Conf. Sensitivity Anal. Model Output* (Procedia Social and Behavioral Sciences, Elsevier, Milan, Italy), 7629–7630.

Chun M-H, Han S-J, Tak N-I (2000) An uncertainty importance measure using a distance metric for the change in a cumulative distribution function. *Reliability Engrg. System Safety* 70(3):313–321.

Clemen RT (1997) *Making Hard Decisions: An Introduction to Decision Analysis* (Duxbury Press, Pacific Grove, CA).

Cox LA (2008) Why risk is not variance: An expository note. *Risk Anal.* 28(4):925–928.

Crnkovic C, Drachman J (1996) Quality control. *RISK* 9(9):139–143.

Deza MM, Deza E (2009) *Encyclopedia of Distances* (Springer-Verlag, Berlin).

Efron B, Gong G (1983) A leisurely look at the bootstrap, the jackknife, and cross-validation. *Amer. Statistician* 37(1):36–48.

Esipenko VI, Shchuko OB (1995) Probability density of the sum of random variables with random coefficients. *Radiophysics Quantum Electronics* 38(6):390–394.

Felli JC, Hazen GB (1998) Sensitivity analysis and the expected value of perfect information. *Medical Decision Making* 18(1): 95–109.

Felli JC, Hazen GB (2004) Javelin diagrams: A graphical tool for probabilistic sensitivity analysis. *Decision Anal.* 1(2):93–107.

Grubbström RW (1967) On the application of the Laplace transform to certain economic problems. *Management Sci.* 13(7):558–567.

Hazen G (2009) An extension of the internal rate of return to stochastic cash flows. *Management Sci.* 55(6):1030–1034.

Hazen GB, Huang M (2006) Parametric sensitivity analysis using large-sample approximate Bayesian posterior distributions. *Decision Anal.* 3(4):208–219.

Hertz DB (1964) Risk analysis in capital investment. *Harvard Bus. Rev.* 42(1):95–106.

Hillier FS (1963) The derivation of probabilistic information for the evaluation of risky investments. *Management Sci.* 9(3):443–457.

Howard RA (1966) Information value theory. *IEEE Trans. Systems Sci. Cybernetics* 2(1):22–26.

Howard RA (1983) The evolution of decision analysis. Howard RA, Matheson JE, eds. *The Principles and Applications of Decision Analysis* (Strategic Decisions Group, Menlo Park, CA), 5–16.

Howard RA (1988) Decision analysis: Practice and promise. *Management Sci.* 34(6):679–695.

Iman RL, Hora SC (1990) A robust measure of uncertainty importance for use in fault tree system analysis. *Risk Anal.* 10(3):401–406.

Jacquet P, Bruner R, Bodily SE (1999) Genzyme/GelTex Pharmaceuticals joint venture. Case UVA-F-1254, Darden School of Business, University of Viriginia, Charlottesville.

Jose VRR, Nau RF, Winkler RL (2008) Scoring rules, generalized entropy, and utility maximization. *Oper. Res.* 56(5):1146–1157.

Jose VRR, Nau RF, Winkler RL (2009) Sensitivity to distance and baseline distributions in forecast evaluation. *Management Sci.* 55(4):582–590.

Kreps DM (1988) *Notes on the Theory of Choice*, Underground Classics in Economics (WestView Press, Boulder, CO).

Kuiper NH (1960) Tests concerning random points on a circle. *Proc. K. Ned. Akad. Wet. Ser. A* 63:38–47.

Lopez JA, Saidenberg MR (2000) Evaluating credit risk models. *J. Banking Finance* 24(1–2):151–165.

Mason DM, Schuenemeyer JH (1983) A modified Kolmogorov-Smirnov test sensitive to tail alternatives. *Ann. Statist.* 11(3):933–946.

Oakley JE, O'Hagan A (2004) Probabilistic sensitivity analysis of complex models: A Bayesian approach. *J. Royal Statist. Soc. Ser. B* 66(3):751–769.

Oakley JE (2009) Decision-theoretic sensitivity analysis for complex computer models. *Technometrics* 51(2):121–129.

Owen AB (2003) The dimension distribution and quadrature test functions. *Statistica Sinica* 13(1):1–17.

Park CK, Ahn KI (1994) A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment. *Reliability Engrg. System Safety* 46(3): 253–261.

Pearson K (1905) *On the General Theory of Skew Correlation and Non-Linear Regression*: *Mathematical Contributions to the Theory of Evolution—XIV*, Drapers' Company Research Memoirs, Biometric Series II (Dulau & Co., London).

Plischke E (2010) An effective algorithm for computing global sensitivity indices (EASI). *Reliability Engrg. System Safety* 95(4): 354–360.

Plischke E, Borgonovo E, Smith CL (2013) Global sensitivity measures from given data. *Eur. J. Oper. Res.* 226(3):536–550.

Pratt JW, Raiffa H, Schlaifer R (1995) *Statistical Decision Theory*, 3rd ed. (MIT Press, Cambridge, MA).

Ramsay CM (2008) The distribution of sums of i.i.d. Pareto random variables with arbitrary shape parameter. *Comm. Statist., Theory Methods* 37(14):2177–2184.

Saltelli A, Marivoet J (1990) Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques. *Reliability Engrg. System Safety* 28(2):229–253.

Savage LJ (1954) *Foundations of Statistics*, 2nd ed. (Dover Publications, New York).

Smith JE (1998) Evaluating income streams: A decision analysis approach. *Management Sci.* 44(12, Part 1):1690–1708.

Wagner HM (1995) Global sensitivity analysis. *Oper. Res.* 43(6): 948–969.

Watson GS (1961) Goodness of fit test on a circle. *Biometrika* 48(1–2):109–114.