# THE ECONOMICS OF MEASURING QUALITY OF LIFE BY THE STANDARD GAMBLE METHOD

**Eric EISENSTAT**, Lecturer Ph.D.
**Manuela EPURE**, Assoc. Prof. Ph.D.
Faculty of Marketing and International Affaires
*Spiru Haret* University
**Patrick Francis GRAY**, Prof. Ph.D.
University of Ulster, UK

***Abstract***

*Of the countless methods for measuring the quality of life (QOL) that have been proposed and analyzed by the vast, multi-disciplinary literature over the past decades, only a select few are alluded to in Romanian scientific circles, and even fewer are actually implemented in practical applications. In adapting existing techniques and engineering a viable measurement system specific to Romania, there exists an important opportunity to re-evaluate the merits and shortcomings of established approaches. As a specific example, we consider the Standard Gamble (SG) method of eliciting "subjective utilities" related to particular health states. The re-evaluation is accomplished from three perspectives: theoretic consistency, interpretability and practicality. It is demonstrated that consistent with economic theory, the appropriate interpretation of SG derived measures is that of a Hicksian change in welfare valuation, rather than a cardinal measure of preferences. A practical consequence of the latter is that SG will necessarily produce a higher QOL value for individuals exhibiting more risk aversion. This leads us to contemplate that in fact SG may be a more appropriate methodology in other contexts, not necessarily health-related, where quality of life may indeed be correlated with the willingness to take risks.*

**Key-words:** *quality of life, measurement, cost-effectiveness analysis, von Neumann-Morgenstern utilities, Romanian well-being index*

**JEL Classification:** $B_{41}$, $D_{81}$, $I_{19}$

## Introduction

The main purpose of this article is to contribute to the methodological development of instruments for the purpose of measuring *quality of life* (QOL) in Romania. By QOL measurement, we refer to the instruments, techniques and scales resulting in a unified quantitative assessment of well-being. Of course, developing methodology that can successfully measure through a minimal set of indicators something as vast and multifaceted as well-being, is indeed a daunting task. Nevertheless, such methods by now boast a long and rich history of evolution in the scientific and public administration circles.

The remarkable development in this field has undoubtedly been sustained by the persevering interest from a public policy perspective. This is hardly surprising – it is not difficult to imagine how a manageable set of quantities that convey information regarding the well-being of individuals or social groups provides an indispensable tool for guiding and monitoring policy. In fact, the interest has recently reached such a significance for policymakers as to compel French President Nicolas Sarkozy to organize a Commission for the purpose of constructing alternative indicators to replace the (typically GDP-based) figures derived from economic data, that would be more suitable for assessing societal well-being, as well as economic, environmental, and social sustainability (Stiglitz et al., 2009).

While the collection of techniques that have been accumulated over the years is far too extensive to collate in general, several noticeable observations stand out. First, QOL measurement methods may be broadly classified as either representing *objective* social indicators or targeting the assessment of *subjective well-being* (SWB) (Diener and Suh, 1997). The latter approach, which accounts for individuals' subjective experiences, has gained particular notoriety in recent years. Beyond this classification, the development of specific measurement instruments appears to have taken a context-specific path. Given the vagueness and complexity of deriving a common, exhaustive definition of QOL, it seems natural that the various methodological evolutions have adhered to narrower domains where a more concrete understanding of the QOL construct is possible, and the resulting measures are destined for similar, well-defined applications.

To that end, the most prolific work on such methodology is perhaps in the contexts of social indicators analysis and health-care applications. Stiglitz et al. (2009) and Diener and Suh (1997) provide a comprehensive list of references to the relevant social indicators literature. The line of work related to health-care is the primary focus of this article and is discussed extensively in what follows. However, it is interesting to further note that the development and application of QOL measurement methodologies has also extended to other domains. For example, these techniques have been recently utilized in research related to urban planning (Rogerson, 1999; Larson, 2010; Tazebay et al., 2010) as well as to analyzing livelihood under war-like conditions (Giacaman et al., 2007). In all, it is fair to conclude that regardless of the context, the effort is generally quite inter-disciplinary.

In contrast, the application and development of such methodology to address the needs specific to Romania's society has staggered considerably. Only a few scattered implementations of known techniques may be identified within the Romanian academic literature. These are almost exclusively restricted to the health-care domain, where a limited set of psychometric methods is considered in assessing QOL related to dental services (Lupu, 2006; Campian et al., 2008), kidney replacement (Sinescu et al., 2008) and mental health (Ardelean et al., 2009). Public administration use of such tools is virtually nonexistent. Given this state of development, we focus our attention on health-care, where QOL has acquired the nomenclature, *health-related quality of life* (HRQoL).

Nevertheless, our principle thesis is not restricted to the health-care domain, whereas the message we wish to advocate is the following. It is undisputable that the public policy interest in such tools, as discussed above, not only extends to Romania, but is in fact magnified as such given the rough period of transition and growth that currently prevails. Therefore, the question is not *whether* the development of QOL measurement methods, customized to the particularities of the Romanian society, is a worthwhile endeavor. Rather, what is of interest is *how* this development should be approached. To that end, ignoring the existing work on the relevant measurement methodology would be clearly inefficient – *adapting* existing instruments to the Romanian context is undoubtedly a more effective approach than reinventing such methods from scratch. On the other hand, blindly *adopting* existing measurements is likewise unadvisable.

In addressing the latter, our main hypothesis is that inherent to the process of adapting existing techniques and engineering a viable measurement system specific to Romania, there exists an important opportunity to re-evaluate the merits and shortcomings of established approaches. Moreover, this re-evaluation need not be viewed strictly in a negative light. As a specific example within health-care, we consider the *standard gamble* (SG) method of eliciting "subjective utilities" related to particular states of health, in generating direct measures of HRQoL. Such measures have been applied in numerous medical contexts over the past four decades in conjunction with *cost-effectiveness analysis* (CEA) of various health-care related programs.

Consequently, we approach the re-evaluation from three angles. First, we re-examine the *theoretical* basis that founds the generally accepted justification for using SG in health-care. Based on our findings, we re-derive a more appropriate *interpretation* of the measures produced by the SG method as numerical assessments of QOL. Finally, taking together the theoretical and interpretative aspects, we re-evaluate the practicality of the SG method in health-care applications. In large, we find that SG-derived values may be suitable in certain contexts, but should be approached with caution. In this sense, the analysis suggests that such measures may be in fact more fitting in assessing QOL outside of health-care, in domains where until now such methods have yet to be considered.

## Health-Related Quality of Life Measurement

Quantitative assessment of an individual's well-being related to health can be traced back to at least the work of Karnofsky and Burchenal (1949). Since then countless scales, techniques, procedures, etc. have been invented, debated, put into practice as well as rendered obsolete. The main aim of such assessments is to guide decisions regarding health-care initiatives, where a central goal is often improving patients' quality of life. In this context, existing methodologies may be classified in terms of four major categories: *psychometric scales*, *quality of well-being* scales (QWB), *health utility indices* (HUI) and *utility measurements*.

All methods related to HRQoL assessment are essentially subjective measures. Classic psychometric scales in this context are derived based on information provided by individuals regarding the various dimensions of their health states. In particular, individuals are asked to rate specific aspects of their physical, mental and social functioning. These scores are subsequently combined in a standardized way to produce a unified assessment. In consequence, such psychometric scales encompass two *objective* properties: (i) establishing the dimensions of well-being as well as the specific elements that are relevant for its assessment, and (ii) assigning objective weights to each scored item in arriving at an overall assessment. QWB and HUI add a subjective layer to a typical psychometric scale by modifying (ii) such that the weights assigned to scored items are also provided by the individual. In this respect, QWB/HUI are *hybrid* instruments – forming a bridge between psychometric scales and utility measurements (Revicki and Kaplan, 1993).

In contrast, methods falling within the class of utility measurements treat HRQoL as an *entirely* subjective construct. In this view, the various elements of mental, physical and social functioning define an individual's *holistic* health state. For each individual, in turn, various health states convey different degrees of appreciation or value, commonly referred to as the *utility of the health state*. Hence, utility measurements focus on assessing this final result – the subjective valuation of a particular holistic state of health. Implicit in such procedures is the subjective determination of what the various dimensions consist of, their degree of contribution to the overall well-being, as well as the way in which specific elements of the health states are combined in determining the final valuation. In other words, the individual decides the functional form of the utility as well as input variables; the researcher only observes the outcome of the subjective assessment.

The SG method belongs to the utility assessment category of HRQoL measures. More specifically, it is part of a small set of techniques that generate *elicited utilities*, which is a notably different approach in that it does not require individuals to explicitly rate any aspect of their well-being. Instead, participants are asked to complete certain tasks that *reveal* implicitly the individual's valuation of a given health state. The SG method, along with its close sibling the *time tradeoff* (TTO) method were originally proposed by Torrance et al. (1972) and further popularized by Torrance (1986, 1987).[1] According to Torrance (1986), the methods find their motivation in the economic concept, *willingness-to-pay* (WTP). The

[1] To be clear, while the TTO method is an entirely novel technique attributable to Torrance et al. (1972), the SG method is a direct adaptation of (one variant of) a technique by the same name founded in decision theory (see for example, Farquhar, 1984). The originality, therefore, lies in its application to measuring HRQoL.

latter, it is argued, is particularly suitable for economic appraisal of medical initiatives and CEA. However, its use in the medical field has been generally discouraged by operational difficulties.

Accordingly, SG and TTO provide alternative methods of the same nature as WTP, but with the advantage that they are simpler to implement in practice. Hence, their application facilitates CEA, or more appropriately in this case, *cost utility analysis* CUA. Interestingly, the original conception of the methods establishes SG as the theoretically founded methodology, with TTO being a more practical alternative (in that participants find it easier to relate to TTO tasks, compared to SG tasks), albeit lacking any theoretical underpinning. What is the claimed theoretical justification for the SG technique? Simply, the argument unveils, it yields a *cardinal utility* measure of the individual's *strength of preferences* for a particular health state, which is in direct alignment with modern economic utility theory.

Incidentally, over two decades after its initial appearance, the TTO method was endowed with a theoretical justification by Buckingham and Devlin (2006), who elegantly demonstrated that TTO may be correctly interpreted as a Hicksian change in welfare measure. This interpretation, of course, bears a strong resemblance to WTP: while WTP measures welfare change by quantifying (for example) how much wealth/income an individual would be willing to forego to achieve a better health-state, TTO provides a corresponding measure in time units by quantifying how much (life) time the individual would renounce to achieve the same health improvement. Therefore, TTO is indeed firmly grounded in economic theory. Meanwhile, the cardinal utility explanation supporting SG-derived values has largely prevailed – that its interpretation as a measure of cardinal utility validates the SG method on a theoretical level is systematically accepted by the vast literature that it has inspired.

To an economist, this is of course unsettling. The concept of cardinal utility has been widely abandoned by the majority of economic analysis in favor of the more methodologically sound concept of ordinal utility. While a formal analysis is beyond the scope of the present discussion (interested readers are referred to Fishburn (1989) for an eloquent account of the history of utility theory in economics), we briefly point out that central to this view are two aspects: (i) the axiomization required to establish cardinality in utility has proven to entail formidable problems in relating it to credible consumer psychology, and (ii) cardinality is largely unnecessary for economic analysis in the sense that an ordinal ranking of preferences is sufficient to rigorously establish economic results. From an economics perspective, utility plays little more than the role of an operationally convenient tool – it is a quantitative index representing preferences that greatly simplifies analytical operation (e.g. it is far simpler to maximize a utility function than to search for the most preferred bundle of goods among all possible bundles).

How does the medical care literature arrive at the conclusion that the SG method produces a cardinal measure of an individual’s preferences? In essence, by ignoring the first point and dismissing the second as irrelevant. For example, Torrance (1987) proclaims that “a simple definition of utility is that it is a cardinal measure of the strength of one’s preference.” This declaration is further sustained several explicit assertions that while ordinal utility is conventional for microeconomists, an equally valid view of utility is that of a cardinal valuation of one’s strength of preferences; all that is necessary is a simple switch of axioms. In this way, utility theory is easily extendible outside of the microecnomic context, and all the relevant precautions – such as the impossibility of aggregating utilities across individuals – s may be readily ignored.

Unfortunately, the SG method is *not* vested with a proprietary set of axioms that formally establish the claimed cardinality of preferences. Rather, this complication is circumvented by appealing to the axioms of von Neumann and Morgenstern (vN-M), upon which their utility theory is founded (vonNeumann and Morgenstern, 1944). From there, the cardinality is forced directly by tying vN-M utilities to cardinal measures of strength of preferences. Unfortunately, within the context of economic theory, which defines the setting for von Neumann and Morgenstern’s work, no such connection is possible, a fact that is firmly rooted in the economics perspective since at least the 1950’s. Section 3 is dedicated to a more in-depth examination of why this is so.

From a practical perspective, despite their overall popularity, utility elicitation methods have encountered a number of conceptual and methodological difficulties in health-care applications. For example, they sparked considerable controversy regarding the use of utility assessments as a basis for administrative decision making in the medical care practice (Mulley, 1989). On the methodological side, attention is devoted to the *invariance* and *inconsistency* confounding HRQoL measures derived from utility elicitation, where invariance refers to participants exhibiting an equivalent preference for distinct health states, and inconsistency describes situations where individuals yield higher valuations for *clinically* inferior health states. Interestingly, both types of errors seem to be more prevalent in SG implementations.

In this context, the errors are generally attributed to the inability of participants to correctly complete the task at hand. From this premise, focus has been mainly concentrated on either linking participant incompetence to personal characteristics such as education level, ethnicity, etc. (Bravata et al., 2010), or on attempts to design procedures to “correct” for the incompetence (Lenert et al., 2003). The latter mainly involves intercepting “illogical” responses and persuading participants to modify them. Evidently, what constitutes an illogical response is determined by the researcher, presumably based on the *objective* clinical evaluations of the health states. Of course, one might wonder as to the purpose of soliciting purely subjective assessments of well-being, which are required to conform to some predetermined objective evaluative criteria.

Perhaps of the more serious methodological issues identified, is the internal inconsistency of the SG method. In particular, Llewellyn-Thomas et al. (1982) through a series of experiments find that changes in gamble outcomes exert a significant influence on the resulting valuations. Although it has not received as much attention as the invariance/ inconsistency topic, this arguably presents a far greater cause for concern. In fact, it is by now well known in the decision analysis literature (where the SG method finds its roots) that this and similar types of preference elicitation methods embody a variety of important *response biases*. Accordingly, a comprehensive survey and analysis of this subject is provided in Schwand et al. (2010), most of which has yet to find its way into the relevant health-care literature.

Related to the methodological concerns uncovered and omitted by the health-care community, we wish to emphasize the following. To properly identify, interpret and possibly correct sources of errors in utility elicitation procedures, it imperative to unchain this methodology from the burden of cardinality. For example, as demonstrated in subsequent sections, invariance in many situations arises quite naturally in the SG context, where only minor errors on the part of the participants are sufficient to generate indifference in revealed preferences. This serves to underline precisely a central theme of this article. The purpose of re-evaluating the theoretical justification of a particular method need not necessarily be to invalidate it. Rather, the more positive motivation as in our case, may be to construct a more accurate interpretation, and in doing so, to allow for a more effective assessment of how and where the methods are best applied, as well as to guide extensions that better account for the various sources of related errors.

### The Standard Gamble and Expected Utility Theory

The standard gamble methodology is based on a remarkably simple procedure. In short, the SG task lays out to the participant two alternatives. Under the first alternative, the individual would remain in a particular state of health (guaranteed) for the remainder of her life. The second alternative, however, involves a gamble with two possible outcomes; with probability $\pi$ the participant would attain perfect health for the rest of her life, while with probability $1-\pi$ instant death would result. The objective, then, is to find the probability $\pi_*$ at which the individual is exactly indifferent between alternatives $1$ and $2$.[2]

To understand how this procedure is related to the vN-M expected utility theory, let us restate a (grossly oversimplified) rendition of the theory's main result, as is fit for our purposes. To this end, consider some possible outcomes $y_l, y_m, y'_m, y_u$ that can be ordered in terms of preferences such that either $y_m \prec y'_m$ (e.g. $y'_m$ is preferred to $y_m$), $y_m \succ y'_m$ (e.g. $y_m$ is preferred to $y'_m$) or $y_m \sim y'_m$ (e.g. $y_m$ is indifferent to $y'_m$), with a presupposed ordering $y_l \precsim y_m \prec y_u$, $y_l \prec y'_m \prec y_u$. Under some “reasonable axioms,” for any lottery $L$ which results in $y_u$ with probability $\pi$ and $y_l$ with probability $1-\pi$ (denoted by

[2] For more details, especially those pertaining to health state description, see Torrance (1986, 1987).

the triplet $L \equiv (\pi; y_l, y_u)$ ), and similarly $L'$ which results in $y_u$ with probability $\pi'$ and $y_l$ with probability $1-\pi'$ (denoted as $L' \equiv (\pi'; y_l, y_u)$ ), the axioms hold *if and only if* there exists a *real valued* function $v$ defined by

$$v(y_m) = v(L) = \pi v(y_u) + (1-\pi)v(y_l) \tag{1}$$

such that

$$L \succ L' \Leftrightarrow v(L) > v(L') \tag{2}$$

Hence, directly employing the specification (1) and arbitrarily setting $v(y_l) = 0$ and $v(y_u) = 1$, it is clear that the probability that defines the lottery determines the numerical utility value as $v(h_m) = \pi$ .

Mathematically, the SG method is predicated on the foregoing conclusion. The procedure elicits "subjective utility values" by requesting an individual to specify the probability $\pi$ that establishes indifference between the certain outcome $y_m$ and the lottery $L$. Consequently, with $\pi$ and $\pi'$ obtained in this manner, a comparison of the resulting numerical values corresponds directly to the comparison in preferences between the outcomes $y_m$ and $y'_m$; the individual prefers $y_m$ to $y'_m$ if and only if $\pi > \pi'$ is chosen. However, this the extent of expected utility theory's involvement in the SG methodology; the idea that $\pi$ provides a numerical assessment of an individual's valuation of a particular health state, a value that is moreover comparable across individuals, is *conceptually* predicated on elements beyond the scope of the vN-M theory.

Specifically, two crucial components are incorporated by the SG methodology in extending the vN-M utility index to a quality of life measure:

1) $v(y_m)$ is a *cardinal utility measure* in the sense that it measures an individual's *strength of preferences*;
2) the outcomes (e.g. $y_l$, $y_m$, $y'_m$, $y_u$, etc.) on which expected utility theory operates may be reasonably treated as *health states*.

The first conceptual component is claimed explicitly in the literature that promotes the theoretical validity of the method (Torrance, 1986, 1987). The second is implicit in nature. Both components are essential to the justification of SG-elicited utilities as measuring QOL and neither is trivial. We presently examine each of the components in further detail.

Let us begin the analysis of the cardinality claim with an closer reading of Torrance (1986), where it is established that

> the standard gamble measurement technique is valid by definition because it is based directly on the [vN-M] axioms, and the validity of the other techniques can be determined by comparison. (Torrance, 1986, p. 27)

Accordingly, the validity claim arises from the premise that vN-M utilities are themselves cardinal:

> The standard gamble is the classical method of measuring cardinal preferences. It is based directly on the fundamental axioms of utility theory, first presented by von Neumann and Morgenstern. (Torrance, 1986, p. 20)

Astonishingly, the fact this premise is strictly incompatible with economic theory is entirely overlooked.

Since as early as the 1950's, economists have exhaustively demonstrated that the cardinality of the function $v(y_m)$ is strictly a mathematical property. Work by the likes of Ellsberg (1954) and Baumol (1958) were among the first to establish that this property does not bear and cannot be misconstrued as bearing any relation to the economic concept of cardinality in terms of measuring pleasure, appreciation, satisfaction, etc. Likewise, Luce and Raiffa (1957) formally outlined the fallacies of associating vN-M utilities with an assessment of the strength of an individual's preferences.

In fact, in their seminal work von Neumann and Morgenstern never sought such a conclusion. Their undertaking was motivated by a need to operationally simplify the analytical burden pertaining to game theory analysis. Indeed, their purpose was accomplished with irrefutable success; as Baumol (1958) explicitly demonstrates, maximizing vN-M utilities invariably leads to correct predictions of an individual's choice of lotteries. The direct implication of this, of course, is that behavior under uncertainty may be accurately predicted (to the extent that the vN-M axioms hold) by modeling the individual as maximizing expected utility. However, beyond facilitating prediction of individual behavior, there is no other explicit or implicit purpose envisioned by von Neumann and Morgenstern in developing their remarkable theory. Especially, the fact that in their own exposition (von Neumann and Morgenstern, 1944) they explicitly divorce their approach from any concept related to a cardinal measurement of preferences, is attested to by the thorough introspective of Fishburn (1989).

We dedicate the remainder of this section to illustrating the economics interpretation of vN-M utilities. Hence, begin by considering a standard textbook example of consumer behavior: exactly two goods are available for consumption, quantities of which $(x_1, x_2)$ are optimally chosen by maximizing utility subject to a linear budget constraint. Assume for the simplicity of the example that utility is of the *Cobb Douglas* form and the budget constraint consists of income $y$ and the price of good $1$, denoted by $p$.[3] By convention, good $2$ is treated as the

[3] Our main results, however, are readily extendible to the general utility function form $u(\cdot)$.

*numeraire* good with unitary price. Consequently, the mathematical formulation is typically expressed

$$\max_{x_1, x_2} u(x_1, x_2) \quad \text{s.t.} \quad x_2 \leq y - px_1 \qquad (3)$$

Assuming the consumer is completely informed (e.g. all *exogenous* information is known with certainty), it is well known that setting either $u(x_1, x_2) = x_1^{\alpha} x_2^{\beta}$ or $u(x_1, x_2) = \alpha \ln x_1 + \beta \ln x_2$ leads to the same optimal choices

$$x_1^* = \frac{\alpha}{\alpha + \beta}(y/p) \qquad (4)$$

$$x_2^* = \frac{\beta}{\alpha + \beta} y \qquad (5)$$

In fact, the very same combination of good $1$ and good $2$ is chosen under *any* monotonic transformation of $u$. It is in this sense that the utility function $u(x_1, x_2)$ is regarded as ordinal.

Continuing with the example under certainty, substituting the optimal consumption values $x_1^*, x_2^*$ into the utility function yields the *indirect* utility function

$$v(p, y) = \frac{\alpha^{\alpha} \beta^{\beta}}{(\alpha + \beta)^{\alpha+\beta}} \frac{y^{\alpha+\beta}}{p^{\alpha}} \qquad (6)$$

Note that this expression involves only the exogenously determined price and income. Intuitively, it is referred to as indirect utility because the consumer does not derive pleasure or satisfaction directly from price and income. Rather these variables have a direct effect on an individual's choice of consumption bundle, which in turn, directly determines her utility.

The significance of the latter point is the following. If the price of good $1$ decreases exogenously, it is typically regarded as having a positive effect on the individual's utility. However, this is not due to a direct relationship as the individual does not care about the decrease in price, *per se*. On the other hand, assuming that $y$ remains unchanged, the lower price level effectively expands her budget constraint such as to allow a more favorable choice of $x_1, x_2$. Of course, consuming a more preferred combination of the two goods indeed brings a higher degree of satisfaction to the consumer.

Furthermore, we may ask the question how much income would the individual be willing to give up in order to obtain the price decrease, i.e. to be just as well off with the price decrease (and lower income) as she was originally? Such a question leads us to consider the tradeoff between price and income at a constant utility level. In other words, we utilize the *indifference curve* between price and income derived from the indirect utility function. If the price decrease under consideration is infinitesimal, then the answer is the marginal rate of substitution between price and income, times the price decrease (e.g. $dy = \mathrm{MRS}_{p,y} dp$).[4] Alternatively, if we examine a more substantial price decrease, then the necessary income reduction ($\Delta y$) is referred to as the *compensating variation* of the price decrease, and is computed from the indifference curve as $\Delta y/y = 1-(1-\Delta p/p)^{\frac{\alpha}{\alpha+\beta}}$. Neither $dy$ nor $\Delta y$ is affected by monotonic transformations of $u$.

It is also worthwhile to observe that if the original utility level is denoted as $u$ and the utility achieved under the decreased price as $u'$, then $\Delta y$ measures the *distance* between utility levels $u$ and $u'$. Two comments regarding this observation are in order. First, the metric, in this case, is clearly specified – utility changes are being measured in *monetary* units. Second, while the income-based metric is obviously cardinal (it is defined up to a fixed origin and a scale in terms of the particular monetary unit, e.g. dollar, euro, etc.), the utilities whose distance is being evaluated are still ordinal. That is, the same numeric distance results regardless of the particular monotonic form ascribed to $u$ and $u'$. As will be seen shortly, the latter bears a strong tie to the interpretation of the vN-M metric as a cardinal measure.

Expanding on our simple example, let us consider the individual's choice when the income level is no longer known with certainty, but rather the only information available is

$$y = \begin{cases} y_u & \text{w.p.} \quad \pi \\ y_l & \text{w.p.} \quad 1-\pi \end{cases} \tag{7}$$

Thus, we wish to analyze the situation where a consumer must choose the optimal consumption bundle while facing a lottery as the source of her budget constraint, instead of a fixed income level. Since the budget constraint is binding, we may assume without loss of generality that the consumer will explicitly choose a specific consumption level of good $1$, and let the quantity of good $2$ be determined *ex post* by $x_1$ as well as the outcome of the income lottery.

Therefore, if the consumer's psychology is consistent with the vN-M axioms, then their expected utility theory tells that the consumer determines her optimal consumption bundle by maximizing expected utility as

$$x_1 = \max_{x_1}\{\pi\tilde{u}(x_1, y_u - px_1) + (1-\pi)\tilde{u}(x_1, y_l - px_1)\}$$

$$x_2 = y - px_1$$

[4] By Roy's Identity, it is of course the Marshallian demand, times the price decrease.

The point of emphasis is that the solution to (8) *does* depend on the functional form of $\tilde{u}$. That is, $\tilde{u}(x_1, x_2) = x_1^{\alpha} x_2^{\beta}$ or $\tilde{u}(x_1, x_2) = \alpha \ln x_1 + \beta \ln x_2$ now lead to different solutions of optimal consumption levels $x_1^*, x_2^*$. It is tempting to conclude from this that under uncertainty, expected utility maximization does indeed induce cardinality in the utility index, i.e. under expected utility maximization the utility function conveys quantifiable information on the strength of preferences.

Such a conclusion, however, is largely misguided. The relevant question is whether observing choice under uncertainty would reveal more information about the consumer's preferences with respect to goods $1$ and $2$. The answer is no. Specifically, letting $\tilde{u}(x_1, x_2) = x_1^{\alpha} x_2^{\beta}$ it can be easily shown that solving (8) is equivalent to maximizing

$$\max_{x_1, x_2} u(x_1, x_2) \quad \text{s.t.} \quad x_2 \le \left[\pi (y_u - px_1)^{\beta} + (1-\pi)(y_l - px_1)^{1-\beta}\right]^{\frac{1}{\beta}}$$

while if $\tilde{u}(x_1, x_2) = \alpha \ln x_1 + \beta \ln x_2$, then the corresponding maximization problem is formulated as

$$\max_{x_1, x_2} u(x_1, x_2) \quad \text{s.t.} \quad x_2 \le (y_u - px_1)^{\pi} (y_l - px_1)^{1-\pi}$$

It can be shown that the above is extendable to any monotonic transformation of $\tilde{u}$: such transformations do not affect the underlying utility function $u(x_1, x_2)$ that captures the individual's attitude towards the consumption of the two goods. Instead, monotonic variations of $\tilde{u}$ in expected utility maximization simply reflect differences in how the individual perceives her *budget constraint* under uncertainty.

The concept is depicted graphically in Figure 1. If either $y_l$ or $y_u$ is known to occur with certainty, then the budget constraint is linear (at the respective income level) and standard textbook analysis applies. However, when the income occurs due to chance, such that deterministic income in the budget constraint is replaced by the lottery $(\pi; y_l, y_u)$, then the *perceived stochastic budget constrained* (PSBC) takes on a curved shape between the two limiting budget constraints. The "perceived" element in the labeling reflects the fact that the particular shape of this budget constraint is subjective – it depends on the individual's attitude regarding risk. For example, risk neutrality in this case may be attributed to a linear PSBC. In general, the shape of the PSBC will be weakly concave, and therefore, interior solutions prevail under uncertainty.

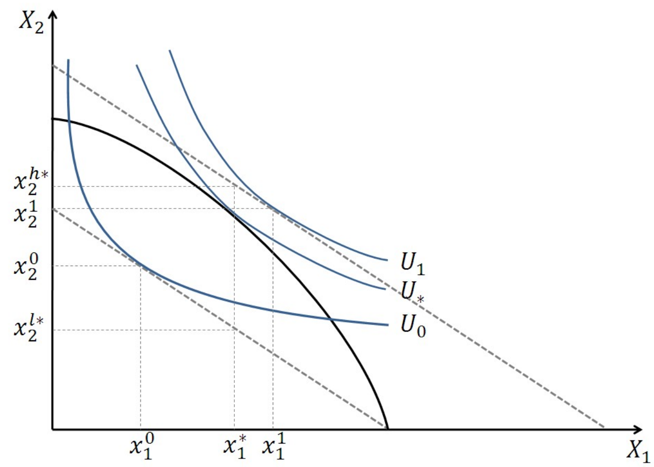


Fig. 1. *Perceived Stochastic Budget Constraint*

It is worthwhile to emphasize that the indifference map remains unchanged regardless of the individual's information about income or her risk attitude. Hence, there is no novelty being generated by uncertainty as to the strength of preferences between bundles of good 1 and good 2. On the other hand, the probability $\pi$ on which the lottery $L \equiv (\pi; y_l, y_u)$ is founded, serves to shift the PSBC similar to the way in which the income level shifts an ordinary budget constraint in the full information case.

The analogy, therefore, is that under uncertainty we replace income $y$ with the lottery $L$. Consequently, indifference curves based on indirect utility functions may be constructed between elements of $L$ and other exogenous factors.[5] If we assume that the the support of the lottery (i.e. $y_l, y_u$) remains fixed, the latter amounts to posing exogenous factors against $\pi$. Accordingly, $\pi$ may be used to measure a change in welfare induced by those alternative exogenous variables, and in consequence, $\pi$ represents a metric by which utility distances may be measured, in the spirit of the income-based metric relevant under certainty. That is, instead of $\Delta y$ we may consider differences in the lotteries, whereby denoting $L_0 \equiv (0; y_l, y_u)$ and $L_* \equiv (\pi_*; y_l, y_u)$, define the difference operator $\Delta_{0,*}$ on $L$ as the difference between $L_0$ and $L_*$, respectively

$$\Delta_{0,*}L \equiv \pi_* \qquad (11)$$

[5] Smidts (1997) refers to vN-M utilities, in interpreting Fishburn (1989), as "probabilistic indifference curves measured by means of lotteries."

This definition is operational to the extent that it is consistent with measuring welfare change between any utility levels represented by indifference curves passing through the area between the limiting budget constraints. For example, following Figure 1, $\Delta_{0,*}L$ (equivalently $\pi_*$) measures the distance from $u_0$ to $u_*$. Moreover, one can measure the distance from $u_*$ to $u_1$ by computing

$$\Delta_{*,\dagger} = \Delta_{0,\dagger}L - \Delta_{0,*}L = \pi_\dagger - \pi_* \qquad (12)$$

The important question then becomes what is the nature of such a metric? That is, if income quantifies differences in utilities on the money-metric scale, then what is the unit of measurement induced by probabilities? One possible interpretation is that $\pi$ measures utility differences in *risk* units. Section 4 examines this idea further, in terms of health states. A topic worthy of a more profound analysis, of course, is whether this is reasonable measurement unit for health-care applications.

**Expected Health vs. Risk Tradeoff**

The exposition of the preceding section illustrated that in the standard economic perspective, vN-M utilities bear no connection to the concept of cardinal utility, in the sense of measuring strength of preferences. Instead, it was demonstrated that the numeric value assigned to such a utility by equating it to a pertinently elicited probability, is more appropriately interpreted as a change in welfare measure, in risk units. Observe that the analysis thus far has concentrated on income as outcomes relevant for the vN-M expected utility theory. Income, of course, plays a well defined role in economic analysis and is straightforwardly related to consumer choice. Accordingly, while our main interest is not measuring utility related to income, but rather that related to health states, this conventional framework is useful for at least two reasons: (i) it serves to clarify the theoretical basis for the SG method as applied in QOL assessment, and (ii) it offers guidance as to how health states may be incorporated into the corresponding theory as the relevant outcomes.

To expand on the latter, it is imperative that a distinct, conspicuous connection is established between health states and individual choice. Otherwise, valuing such health states is exclusively a psychology exercise, not an economic one, and involving the vN-M theory in this case is taking it entirely out of context. The manner in which health states may be related to choice deserves serious reflection. If a direct analogy is established between income and health states, then one must keep in mind that income is conventionally considered to affect utility only indirectly by influencing the optimal consumption of a bundle of goods. If, on the other hand, a certain health state is regarded a entering the utility function directly, one must take extra care in correctly interpreting a methodology that departs from the conventional framework.

One particular example of a model that relates health states to choice is that of Meltzer (1997), with extensions particular interesting in the context of the present discussion by Basu et al. (2005, 2009, 2010). In this model, an individual is able to affect her health state at the expense of consumption by investing in healthcare and reducing her disposable income:

$$\max_{m} u\big(h(m), c\big) \quad \text{s.t.} \quad c \le y - m \qquad (13)$$

Therefore, a health states embodies both endogenous (that part that can be affected investment level $m$) and exogenous components. Moreover, the model is easily extendible to choices arising outside of consumption. For example, one might easily replace income with time and consumption with leisure to obtain a tradeoff between health and leisure.[6] In the latter case, an individual would be faced with the choice of how much of her leisure time to dedicate to improving her state of health (e.g. exercising, etc.), where once again she derives utility directly from both elements.

In application to the model in (13), and to a reasonable extent more genetically, the SG method may be related to expected utility theory as follows. When faced with a SG task, a participant in the study will choose the break-even probability $\pi$ by considering the tradeoff between risk and expected health. Specifically, if she prefers more health and less risk, then an indifference map to reflect her decision may be constructed as depicted in Figure 2.

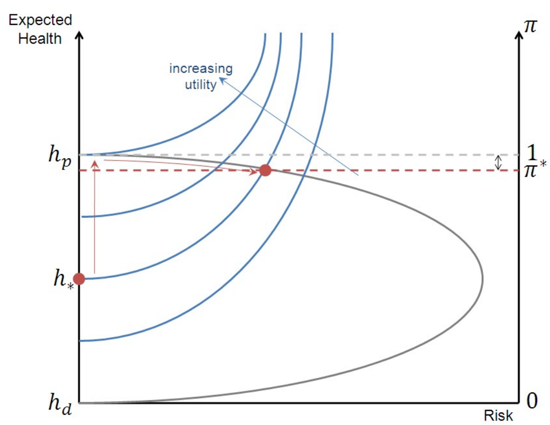


Fig. 2. *Risk vs.Expected Health*

Accordingly, the risk – expected health tradeoff is also directly embodied in the gamble itself, where all feasible combinations are designated by the rotated "U" curve in the plot.

---

[6] We acknowledge that, as was pointed out by participants at the 2011 Symposium "Dreptul la bunăstare – viitorul economiei româneşti", time and income are not generally interchangeable since income is feasibly infinite while time is necessarily finite. In some special cases, this certainly is an important consideration. Nevertheless, to the extent that in a significant portion of economic problems income is treated as a finite constraint on the individual's choice, the extension to time/leisure (and the like) seems reasonably appropriate.

Furthermore, the shape of the indifference curves, in this case, directly relates to the degree of *risk aversion* characterizing the participant in the SG task. Specifically, more risk averse participants are represented by more convex, steeper sloping indifference curves (accordingly, flat indifference curves imply risk neutrality). Moreover, each indifference curve relates a certainty equivalent health state (e.g. with zero risk) to a lottery with a higher expected health and greater risk. The fact that the two lie on the same indifference curve, obviously indicates that the participant is indifferent between the two options. Therefore, we may envision measuring the distance between the utility of some health state $h_*$ and that of perfect health $h_p$, by moving away from the riskless point $h_p$ along the inverted "U" curve, until the indifference curve capturing $h_*$ is intersected. At this intersection, lies the lottery defined by $\pi_*$, which provides the same amount of utility to the participant as the riskless health state $h_*$. Consequently, the distance between utilities derived from $h_*$ and $h_p$ is quantified as $1-\pi_*$.

Hence, this is another way to illustrate the fact that the probabilities elicited by the SG method, in fact, measure distances between utilities (or changes in welfare) rather than any cardinal utility values. More importantly, by considering the measurement procedure in this light, we explicitly account for the crucial element driving the resulting valuation – the participant's risk aversion with respect to the SG task. An important consequence of this is that it offers one potential explanation of the invariance often found in SG-based quality of life measures: a higher degree of risk aversion and hence steeper indifference curves, lead to a greater proportion of health states to be associated with lotteries concentrated at similar probability values (close to 1), where distinctions are more difficult for participants to identify. Therefore, risk aversion may be one simple reason for the encountered invariance in practice.

**Conclusions**

On theoretical grounds, the probabilities elicited by the SG method cannot be justified as measuring a cardinal utility index that quantifies the individual's strength of preferences; ascribing such an interpretation of cardinality to vN-M utilities is at odds with conventional economic theory. On the other hand, because vN-M utilities may be appropriately viewed as measuring welfare change or utility differences, as discussed in section 3, the QOL measure generated by the SG method may be interpreted form this perspective as well. The latter is, of course, predicated on successfully relating health states to an individual's choice, such as to make economic decision theory applicable.

In this respect, economic theory suggests that SG-elicited probabilities may be regarded as capturing utility differences quantified by probabilistic risk units. This interpretation aligns SG with the closely related utility elicitation methods TTO and WTP as measures of welfare change. However, while the QOL index generated by TTO is readily interpretable as *quality adjusted life years* (QALYs) (Torrance et al., 1989), and likewise for the WTP method as *quality adjusted wealth* (QAW), the corresponding QOL index derived from the SG method does not convey a similarly obvious, natural interpretation.

The most concrete characterization of SG is that it reflects the individual's attitude towards risk. The question, however, is whether this is an acceptable property of a QOL index? According to Lenert and Kaplan (2000), it is in fact desirable. Yet, they offer little explanation as to why this should be so, which leaves a number of important gaps in the analysis. The fundamental problem is relating risk attitude to an individual's valuation of health states. Why should a particular health state's effect on an individual's well-being be driven by her attitude towards risk? Lenert and Kaplan (2000) make reference in this regard to a potential decisions one might face when presented with a risky surgery (as a treatment option, say, for a serious illness).

Care must be taken, however, in assimilating the risk attitude exhibited by an individual in completing the SG task to the risk attitude present in deciding upon risky treatments – risk attitudes are likely to take on dramatically different forms across various contexts. In addition, serious illness accompanied by risky treatment options characterizes only a small fraction of situation in which a subjective valuation of health state is typically sought. In general, there is no obvious place for uncertainty in an individual's decisions related to health, and therefore, no clear connection between how risk attitude inherent to the SG task reflects such decisions and the utility derived from them.

To expand on the latter perspective, one might generally inquire whether the risk attitude captured by the SG-elicited probability truly reflects a risk preference of the *individual* or just that of the *participant*. In other words, does a participant's risk attitude, which influences her decisions in completing the SG task, extend beyond the experimental setting? The decision theory literature that Lenert and Kaplan (2000) cite in supporting the desirable risk preference property of the SG method takes into serious consideration the concept of *intrinsic risk attitude* (Bell and Raiffa, 1982; Dyer and Sarin, 1982; Smidts, 1997). In this view, an individual facing a choice under uncertainty is induced with a certain risk preference specific to the situation and not related to any underlying preferences or valuation of resulting outcomes. Consequently, if intrinsic risk attitude is applicable to a particular SG experiment, then the captured risk preference is simply an artifact of the methodology.[7]

To what extent the above considerations are significant when applying the SG method depends on the specific case at hand. The point of emphasis is that by better understanding the theoretical framework in which SG-elicited measurements arise, we can better interpret the results obtained according the context within

[7] This is inline with the findings of Hellinger (1989) that in health-related decisions under uncertainty, "risk attitudes are not absolute but are functions of the parameters in the gamble".

which we operate. To that end, the usefulness of the SG method relative to its utility elicitation counterparts (TTO and WTP), as well as other non-utility based measures, is a practical matter. Regardless of which theory we base the derivation of the method on, as long as its results can be properly interpreted and deemed useful as such, the methodology is likewise useful. The common theme of this article is that the latter is generally context-dependent.

Nevertheless, there *is* a property of the SG method with respect to risk that is generalizable across contexts. As is immediately evident from Figure 2 (and may be verified formally without excessive analytic complexity), the SG-based QOL measures invariably reward risk aversion. Specifically, an individual with a higher degree of risk aversion will always yield a higher elicited $\pi$, and thus, will *necessarily* be assigned a higher QOL value for the same health state as an individual with less risk aversion. In our view, this is a somewhat confounding property of HRQoL assessment, especially in application to cost effectiveness analysis. Besides generating considerable difficulties of aggregating such measures across individuals with varying degrees of risk aversion, the ensuing CEA based on such a measure will yield a more favorable assessment to a certain health-care initiative among groups of individuals with more risk aversion, even if differences in health-care improvements across the groups are negligible.

On the other hand, the interpretation suggests that the SG methodology may be well suited in evaluating subjective well-being in situations inherently characterized by risky choices. One might imagine the potential for such situations as arising, for example, in connection with living environments. In particular, this may be an interesting application in conflict regions where "gambles" involving possible improvements in livelihood at the risk of death are often encountered. Risk aversion, in this case, may indeed be correlated with perceived well-being. More specific to Romania, a potentially suitable application may present itself in terms of evaluating subjective well-being of individuals involved in certain hazardous occupations, a topic that, to the best of our knowledge, has yet to be approached from *any* methodological perspective.

## REFERENCES


- Ardelean, L., C. Lupas, D. Spatarel, and M. Stoleru (2009), Calitatea vieţii la personalul din domeniul sanătăţii mintale. *Revista de Neurologieşi Psihiatrie a Copilului şi Adolescentului din România 12*(4), 54-64.
- Basu, A., W. Dale, A. Elstein, and D. Meltzer (2009), A linear index for predicting joint health-states utilities from single health-states utilities. *Health economics 18*(4), 403-19.
- Basu, A., W. Dale, A. Elstein, and D. Meltzer (2010), A time tradeoff method for eliciting partner's quality of life due to patient's health states in prostate cancer. *Medical Decision Making 30*(3), 355-65.
- Basu, A. and D. Meltzer (2005), Implications of spillover effects within the family for medical cost-effectiveness analysis. *Journal of health economics 24*(4), 751-73.

- Baumol, W. J. (1958), The cardinal utility which is ordinal. *The Economic Journal 68*(272), 665-672.
- Bell, D. E. and H. Raiffa (1982), Marginal value and intrinsic risk aversion. In H. Kunreuther (Ed.), *Risk: A Seminar Series*, Laxenburg, Austria, pp. 325-349. IIASA.
- Bravata, D. M., L. M. Nelson, A. M. Garber, and M. K. Goldstein (2010), Invariance and inconsistency in utility ratings. *Medical Decision Making 25*(2), 158-67.
- Buckingham, K. and N. Devlin (2006), A theoretical framework for TTO valuations of health. *Health economics 15*(10), 1149-54.
- Campian, R. S., C. Amariei, B. Crisan, G. Baciut, M. Baciut, O. Lucaciu, and A. Sava (2008), Calitatea vietii corelata cu sanatatea orala si cu cancerul oral si maxilofacial. Working Paper.
- Diener, E. and E. Suh (1997), Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research 40*(1), 189-216.
- Dyer, J. S. and R. K. Sarin (1982), Relative risk aversion. *Management Science 28*, 875–886.
- Ellsberg, D. (1954), Classic and current notions of “measurable utility". *The Economic Journal 64*(255), 528-56.
- Farquhar, P. H. (1984), Utility assessment methods. *Mangement Science 30*(11), 1283–1300.
- Fishburn, P. C. (1989), Retrospective on the utility theory of von neumann and morgenstern. *Journal of Risk and Uncertainty 2*(2), 127-158.
- Giacaman, R., A. Mataria, V. Nguyen-Gillham, R. A. Safieh, A. Stefanini, and S. Chatterji (2007), Quality of life in the Palestinian context: an inquiry in war-like conditions. *Health policy 81*(1), 68-84.
- Hellinger, F. J. (1989), Expected utility theory and risky choices with health outcomes. *Medical Care 27*(3), 273-279.
- Karnofsky, D. A. and J. H. Burchenal (1949), The clinical evaluation of chemotherapeutic agents in cancer. In C. M. Maclead (Ed.), *Evaluation of chemotherapeutic agents*, New York, pp. 191-205. Columbia University Press.
- Larson, S. (2010), Regional well-being in tropical Queensland, Australia: developing a dissatisfaction index to inform government policy. *Environment and Planning A 42*(12), 2972-2989.
- Lenert, L. A. and R. M. Kaplan (2000), Validity and interpretation of preference-based measures of health-related quality of life. *Medical care 38*(9, Suppliment II), 138-150.
- Lenert, L. A., A. Sturley, and M. Rupnow (2003), Toward Improved Methods for Measurement of Utility: Automated Repair of Errors in Elicitations. *Medical Decision Making 23*(1), 67-75.
- Llewellyn-Thomas, H., H. Sutherland, R. Tibshirani, A. Ciampi, J. Till, and N. Boyd (1982), The Measurement of Patients’ Values in Medicine. *Medical Decision Making 2*(4), 449-462.

- Luce, R. D. and H. Raiffa (1957), *Games and decisions: introduction and critical survey*. New York: Wiley.
- Lupu, I. (2006), *Calitatea vieţii în sănătate: Definitii şi instrumente de evaluare. Calitatea Vieţii 17*(1-2), 73-91.
- Meltzer, D. (1997), Accounting for future costs in medical cost-effectiveness analysis. *Journal of Health Economics 16*(1), 33-64.
- Mulley, A. G. J. (1989), Assessing patients' utilities: Can the ends justify the means? *Medical Decision Making 27* (3, suppliment), 269-281.
- Revicki, D. A. and R. M. Kaplan (1993), Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Quality of Life Research 2*(6), 477-487.
- Rogerson, R. J. (1999), Quality of Life and City Competitiveness. *Urban Studies 36*(5), 969-985.
- Schwand, C., R. Vetschera, and L. Wakolbinger (2010), The influence of probabilities on the response mode bias in utility elicitation. *Theory and Decision 69*, 395-416. 10.1007/s11238-010-9193-8.
- Sinescu, I., M. A. Manu, M. Harza, B. Serbanescu, B. Stefan, V. Cerempei, C. Baston, D. Tacu, E. Kerezsy, C. Bucsa, I. Constantinescu, L. Domnisor, E. Burchiu, R. Manu, and O. Palea (2008), Studiul comparativ al calitatii vietii donatorilor de rinichi (tineri / varstnici si inruditi / neinruditi) folosind testul "sf-36 health survey". *Revista Română de Urologie* (3), 1-9.
- Smidts, A. (1997), The relationship between risk attitude and strength of preferences: A test of intrinsic risk attitude. *Management Science 43*(3), 357-370.
- Stiglitz, J. E., A. Sen, and J.-P. Fitoussi (2009), Report by the Commission on the Measurement of Economic Performance and Social Progress.
- Tazebay, I., M. Somuncu, and N. Akpinar (2010), A quantitative assessment for quality of life : The case of metropolitan Ankara, Turkey. *Journal of Agricultural Research 5*(12), 1360-1372.
- Torrance, G. W. (1986), Measurement of health state utilities for economic appraisal. *Journal of health economics 5*(1), 1-30.
- Torrance, G. W. (1987), Utility approach to measuring health-related quality of life. *Journal of chronic diseases 40*(6), 593-603.
- Torrance, G. W. and D. H. Feeny (1989), Utilities and Quality-Adjusted Life Years. *International Journal of Technology Assessment in Health Care 5*(4), 559-575.
- Torrance, G. W., W. H. Thomas, and D. L. Sackett (1972), A utility maximization model for evaluation of health care programs. *Health Services Research 7*(2), 118-133.
- von Neumann, J. and O. Morgenstern (1944), *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.