NBER WORKING PAPER SERIES

EVALUATION OF MEDICAL TECHNOLOGIES WITH UNCERTAIN BENEFITS

Darius N. Lakdawalla
Charles E. Phelps

Working Paper 26058
http://www.nber.org/papers/w26058

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2019

DL acknowledges support from the National Institute on Aging (1R01AG062277). DL discloses that he owns equity in and serves as a consultant to Precision Medicine Group, which provides consulting services to firms in the biopharmaceutical industry. CP discloses that he consulted for Merck Pharmaceuticals on methods to value healthcare in October 2018. No consulting income was received by either author to support this work. The authors are grateful to Ms. Hanh Nguyen for exceptional assistance with production of graphics. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w26058.ack

Evaluation of Medical Technologies with Uncertain Benefits
Darius N. Lakdawalla and Charles E. Phelps
NBER Working Paper No. 26058
July 2019
JEL No. I1,I13,I18

ABSTRACT

Cost-effectiveness analysis (CEA) remains the de-facto method of choice to evaluate and compare medical interventions. Standard approaches to CEA use the average (mean) outcomes from clinical effectiveness studies such as randomized controlled trials. This paper generalizes standard methods to include uncertainty in clinical outcomes and proposes a generalized version of the quality-adjusted life-year (QALY), referred to as a quality- and risk-adjusted life-year (QRALY). Our approach requires new information from clinical studies – not only means and variances of health outcomes, but also skewness. With that added parameter, this paper shows how Taylor Series expansions of expected utility can account for two distinct effects of uncertainty: the "insurance value" of reducing overall risks to health, and the "value of hope" produced by the presence of positively skewed outcomes. Simulations demonstrate that stochastic terms are particularly important when baseline disease severity is high, and mean treatment effects are low. They also demonstrate that the variance-based term has the greatest importance among the stochastic terms, although skewness- and kurtosis-based terms can be significant in some situations.

Darius N. Lakdawalla
University of Southern California
635 Downey Way, VPD 414-K
Schaeffer Center for Health Policy and Economics
Los Angeles, CA 90089-7273
and NBER
dlakdawa@usc.edu

Charles E. Phelps
University of Rochester
30250 S Highway One
Gualala, CA 95445
charles.phelps@rochester.edu

I. INTRODUCTION

For decades, economists have studied how consumers trade off mortality risk and money [1, 2]. The normative implications of this analysis are particularly helpful in contexts where markets for life-extension investments are incomplete or absent. For example, consumers rarely pay directly for safer transportation infrastructure or environmental standards. In contexts where consumer life expectancy is influenced by the decisions of third parties, the economic theory of mortality risk is often used to calculate how much these third parties should spend on risk-reduction.

These insights have particularly relevance for healthcare. Public health insurers, commercial health insurers, and other third-parties routinely make healthcare spending decisions that influence the length and quality of life for consumers. In these contexts, the economic theory of mortality risk-reduction helps answer the question of how much a social planner should spend on extending life. However, it does not address the question of how to allocate resources across a range of competing interventions, and especially not when these resources are allocated by healthcare payers with potentially inefficient incentives. This question has been addressed by the related theory of cost-effectiveness, of roughly the same vintage as the theory of mortality risk-reduction [3].

Cost-effectiveness has long been used in Britain, Canada, and Australia to determine coverage and reimbursement of new medical technologies by health insurer and to evaluate medical technologies in the US and elsewhere. By recent reports, almost 7300 (and counting)

cost-effectiveness analyses of medical technologies have been published.[1] Our current decade has brought with it increasing interest in cost-effectiveness among insurers in the United States as well. The Institute for Clinical and Economic Review (ICER), a US nonprofit organization, now routinely conducts and releases cost-effectiveness studies for use by American healthcare payers and providers.[2] A recent study suggests that nearly 60% of US payers have relied on or consulted cost-effectiveness analyses in their price negotiations or reimbursement decisions [4]. Moreover, a large Pharmacy Benefit Manager (CVS Caremark) recently proposed to link payment for prescription drugs to their cost-effectiveness by limiting reimbursement to a maximum of $100,000 per estimated Quality Adjusted Life Year (QALY) gained.[3] As cost-effectiveness expands its reach, it becomes increasingly important to ensure that its methods accurately reflect individual and societal preferences.

Traditionally, cost-effectiveness analyses have relied on average health outcomes to assess the value of clinical interventions [5]. Yet, focusing on averages overlooks the role of risk and uncertainty in the effects of medical technologies. This issue has become increasingly salient with the rise of personalized and targeted medical technologies that produce heterogeneous effects across different genotypes. Several recent studies illustrate ways in which abstracting from uncertainty can lead to erroneous inferences about value. Technologies that reduce the variance of overall health outcomes generate substantial value to risk-averse consumers, even holding average outcomes fixed [6]. Effective new treatments for a severe illness reduce risk by limiting the probability of harmful complications. Further, technologies

[1] https://cevr.tuftsmedicalcenter.org/databases/cea-registry, last visited March 22, 2019. This lists reports through 2017.

[2] https://icer-review.org

[3] https://www.forbes.com/sites/joshuacohen/2018/09/20/will-cvs-caremark-make-icer-the-american-nice/#1cb23a8b6173 last visited June 18, 2019.

that increase variance may still be incrementally valuable to risk-averse consumers if they also increase the positive skew in the distribution of health outcomes. An example is the "value of hope" exhibited by cancer patients, many of whom prefer a risky therapy that offers a modest chance of a large long-term survival gain [7, 8]. Motivated by the accumulating empirical evidence, a recent economic task force report from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) called for methods to monetize the value of risk reduction and the value of hope in standard cost-effectiveness frameworks [9].

Two decades ago, Garber and Phelps [10] developed a framework for cost-effectiveness analysis rooted in microeconomic theory. By focusing on treatment interventions with non-stochastic outcomes, their framework showed how to conduct cost-effectiveness analysis based on average treatment effects alone. We seek to generalize that framework by providing empirically tractable methods for conducting cost-effectiveness analysis when consumers are risk-averse over health and health technologies produce variable benefits. Our theoretical analysis yields several insights. First, we identify the "certainty-equivalence ratio" in health improvement that quantifies the value of a risky health intervention in terms of certainty-equivalent health units gained. Second, we quantify how the value of a health intervention falls with variance and rises with positive skewness. In so doing, we bring together into a single framework the earlier identification of "insurance value" from health interventions that lower variance, and "value of hope" from health interventions that offer the chance of positively skewed outcomes. Finally, we develop the concept of the "quality- and risk-adjusted life-year" (QRALY), which generalizes the traditional "quality-adjusted life-year" (QALY) that presumes an absence of uncertainty in health outcomes. Our framework takes as inputs the variance, skewness and (if available) kurtosis in clinical trial or quasi-experimental studies of health

outcomes, along with risk preference parameters already estimated in the economics literature. The earlier Garber and Phelps approach nests as a special case of ours, one in which treatment effects are certain.

The valuation of uncertainty in consumer decision making dates back to the pioneering work of Markowitz [11], Pratt [12] and Arrow [13]. These early studies led to the now-classic "mean-variance" tradeoff in financial economics. Building upon these studies and the work of Kimball on higher-order risk attitudes [14], we develop a framework that incorporates mean, variance, skewness, and kurtosis into the valuation of uncertain health improvements. We primarily study the application to cost-effectiveness analysis in healthcare, but the methods are identical for studying stochastic outcomes of investments in other human capital goods like, vehicle safety, environmental quality, building standards, and any other regulation that affects human health and safety. In addition, the methods could be adapted readily to include any scalar summary of multi-dimensional measures of value such as Multi-Criteria Decision Analysis (MCDA) metrics [15].

Our analysis demonstrates that variance in health outcomes is an economic "bad" when consumers are risk-averse, and it shows how health technologies may increase or decrease variance in outcomes. In addition, we show how positive skewness in treatment effects is an economic "good" when consumers are prudent. More generally, we demonstrate how changes in the skewness and kurtosis of distributions of health outcomes can modify the effects of changes in variance. Our framework illustrates how to introduce these elements without having to build complicated decision trees that might seek to capture the same events. Our methods only require new estimates of treatment effect skewness (and, optionally, kurtosis) in health outcomes for competing health technologies.

Our results have implications for the coverage and reimbursement of new medical technologies by health insurers. Accounting for uncertain health outcomes can better align coverage policies with the underlying utility of risk-averse enrollees. Current practices in cost-effectiveness analysis may over-value technologies with highly variable outcomes, but under-value those with positively skewed outcomes. This in itself is an increasingly important issue. A wide variety of diseases exhibit variable treatment responses [16], including highly prevalent diseases like depression [17], rheumatoid arthritis [18], diabetes [19], and cancer [20]. While continued progress in diagnostics might eventually improve the ability to forecast which patients will respond to a particular drug, there remains a considerable amount of variable treatment response that is unknown before treatments are administered. Value assessment methods need to incorporate practical strategies for quantifying the effects of this uncertainty.

We begin by developing a theoretical framework that accounts for uncertain clinical effects in the valuation of medical technologies. We then illustrate the implications of this framework with a series of numerical examples centering on parameter estimates derived from the economic literature.

II. THEORETICAL FRAMEWORK

We aim for a theoretically grounded framework amenable to empirical analysis. We make several simplifications to achieve this. First, we assume that the consumer displays the same degree of relative risk-aversion over health outcomes as over consumption. While our theoretical derivations do not require this assumption, empirical considerations do. To our knowledge, risk-aversion over health outcomes has not separately been estimated in the

consumer economics literature.[4] From a practical standpoint, this suggests one of two approaches: assume the same risk-aversion over composite consumption and health flows, or assume risk-neutrality over health. We believe that the latter approach is no longer tenable, in light of the empirical evidence on consumer risk preferences over health improvements.

Our analysis provides a specific and rigorous method to evaluate the utility from stochastic (risky) health gains. The key pieces of information needed to compare two technologies are the means, variances and skewness of the distribution of health benefits that they produce, with the potential of adding estimates of kurtosis if available. The analysis relies on proof of the economic legitimacy of using cost-effectiveness analysis (CEA) to guide resource allocation and the existence of an optimal cost-effectiveness cutoff K – the willingness to pay (WTP) for one quality-adjusted life-year (QALY) -- which can also be expressed as a multiplier of income (M). In that formulation, $\frac{K}{M} = \frac{1}{E_{UM}}$, the inverse of the elasticity of utility with respect to income [21]. It represents the opportunity cost in terms of foregone consumption of goods and services that can be purchased with income.

In our model, an individual is sick with probability p and well with probability $(1-p)$. The flow of health in the well state is the scalar H_W. The flow of health in the sick state is $H_S(s) \equiv H_W - s$, where the health reduction in the sick state, $s \in (0, H_W)$, is a random variable with probability density $g_s(s)$. For empirical purposes, suppose that health is measured in quality-adjusted life-years (QALYs) per year (or other period) of life. The theoretical analysis

[4] Cordoba and Ripoll [23] estimate a lifetime expected utility model using Epstein-Zin-Weil preferences that separates the parameters for intertemporal substitution (the inverse of the coefficient of relative risk-aversion) and a "mortality aversion" parameter that measures aversion to the risk of dying. The latter is not the same as risk aversion over an index of the health state, which our model requires. Further, their estimate of mortality aversion is calibrated to estimates of the value of a statistical life (VSL) that have many potential upward biases (see [24] Appendix C1a and C1b).

would be unchanged if health were measured as life-years, disability-free life-years, or some other related construct. In this context, health is measured in natural units, not in utility-equivalents as might be derived in the (seldom-used) standard gamble approach. The period utility function is $u(M,H)$, for composite consumption, M, and health level H. The willingness to pay for a QALY equals $K \equiv \frac{u_H}{u_M}$ – the marginal rate of substitution (MRS) between health and general consumption. Expected utility is given by:

$$EU \equiv p\int_{s\in S} u(M,H_S(s))g_s(s)ds + (1-p)u(M,H_W) \qquad (1)$$

The standard QALYs framework specifies $u(M,H) = v(M)H$, for some strictly concave function v [10]. However, if utility is instead concave in H, the usual implications for QALYs will not obtain. To pursue this point further, observe that the willingness to pay for Δ additional QALYs is given by $\frac{u_H(M,q)}{u_M(M,q)}\Delta$ for an individual at baseline QALY level q. Moreover, in the conventional framework, the value of a statistical life-year spent at QALY level q is $\frac{u(M,q)}{u_M(M,q)}$ [2]. The QALYs framework requires that the willingness to pay for Δ additional QALYs be the same, regardless of whether these are purchased via quality-of-life increases (i.e., a gain of Δ QALYs) or via life-extension (i.e., adding $\frac{\Delta}{q}$ years of life at the QALY level q). Therefore, it must be true that $\frac{u_H(M,q)}{u_M(M,q)}\Delta = \frac{u(M,q)}{u_M(M,q)}\frac{\Delta}{q}$. It is straightforward to show that this expression holds at equality only if utility is linear in health, in which case $u_H(M,q)q = u(M,q)$. However, consumer risk-aversion over health breaks the equivalence between QALYs gained via life-extension and via quality of life improvement. Later, we show that the simplicity of QALY-based resource allocation rules can still be preserved by a more generalized formulation – the quality- and risk-adjusted life-year (QRALY).

III. VALUE ASSESSMENT

We wish to assess the value of the incremental health benefit produced by a new medical technology. In the absence of treatment by this technology, a sick patient experiences the baseline health outcome given above, $H_S(s) \equiv H_W - s$, where H_W is the known level of health in the "well" state, and the reduction in health, s, is a random variable with mean μ_s, variance σ_s^2, Pearson's skewness coefficient γ_{1s}, and Pearson's kurtosis γ_{2s}. Thus, $\mu_S \equiv E(s)$, $\sigma_S^2 \equiv E(s-\mu_S)^2$, and $\gamma_{1S} \equiv \frac{E(s-\mu_S)^3}{\sigma_s^3}$ and $\gamma_{2S} \equiv \frac{E(s-\mu_S)^4}{\sigma_s^4}$. For compactness, define the average health outcome in the sick state as $\bar{H}_S \equiv H_W - \mu_S$. The baseline health outcome, H_S, may represent the benefits from a portfolio of "standard of care" technologies, or the health outcome in an entirely untreated state, depending on the context.

Consider the introduction of a new medical technology (T). A sick patient treated with this technology receives the incremental benefit, β, a random variable with probability density $g_\beta(\beta)$ measuring the benefit of this technology compared to the current standard of care. $\beta \in B$ is distributed with mean μ_B, variance $\sigma_B^2 \equiv E(\beta-\mu_B)^2$, Pearson's skewness coefficient $\gamma_1 \equiv \frac{E(\tau-\mu_B)^3}{\sigma_B^3}$, and Pearson's kurtosis $\gamma_{2B} \equiv \frac{E(\beta-\mu_B)^4}{\sigma_B^4}$. We allow for the possibility of correlation between s (the baseline health outcome) and β (the incremental benefit of the technology). If the two are positively correlated, the new technology is particularly effective for the sickest patients, and vice-versa. Define σ_{S+B}^2, $\gamma_{1(S+B)}$, and $\gamma_{2(S+B)}$, as variance, Pearson's skewness coefficient, and Pearson's kurtosis of $H_S + \beta$, respectively.

The incremental *ex ante* expected utility of the technology T, from the perspective of a consumer whose health state is not yet realized is given by:

$$E(\Delta U(T)) \equiv p\left[\begin{array}{c}\int_{s\in S}\int_{\beta\in B} u\big(M,(H_W-s+\beta)\big)g_\beta(\beta)g_S(\mathrm{s})d\beta\, ds \\ -\int_{s\in S} u(M,H_W-s)g_S(s)ds\end{array}\right] \quad (2)$$

Since the nonlinearity in this problem will pose empirical difficulties, we develop Taylor Series expansions of the two terms in $\Delta U(T)$. Taylor Series approximations have long been used to approximate the costs of risk-bearing. In the finance literature, Markowitz (1952) first proposed the use of second-order Taylor Series expansions to create mean-variance approximations to expected utility [11]. Arrow (1963) and Pratt (1964) independently derived their well-known risk premium equation partially on this basis [12, 13]. In subsequent decades, a substantial literature in finance has grown around the topic of Taylor-Series approximations to the cost of risk-bearing [22].

Two points from the literature are especially noteworthy for our application. First, a Taylor-Series approximation may not converge to the object of its approximation *for every value in its domain*. This is a well-known mathematical result that strikes a cautionary note for the use of Taylor approximations to expected utility [23]. We address this issue below by presenting some sufficient conditions for asymptotic Taylor Series convergence. Moreover, Taylor Series convergence is neither a necessary nor sufficient condition for a Taylor approximation with only a handful of terms to be precise [24]. We address this issue in our empirical calibration exercises, by illustrating how quickly our Taylor Series expansion converges in the context of measuring health improvements. We also provide a guide for practitioners to measure the speed of convergence in their particular applications.

Assume that technology is paid for through health insurance—*ex ante*—before the realization of the health state or, equivalently, in both the healthy and the sick states. Normalizing this Taylor Series expansion by the expected marginal utility of consumption, EU_M,

yields the consumer's *ex ante* willingness to pay for the new technology, which we define as $V(T) = \frac{\Delta U(T)}{EU_M}$.

Our analysis presumes that increases in average health benefits can be summarized in a single scalar μ_B, whereas in real situations, multiple dimensions of value could affect health benefits. In our situation, the proper approach would first summarize the different dimensions of quality into a scalar using some process such as multi-criteria decision analysis [15, 25] or discrete choice experiments [26], with weights supplied (for technology evaluation purposes) by an appropriate representative panel of potential patients, or even by individual patients when considering specific treatment options available to them (e.g., different cancer therapy regimens or surgical alternatives). Then the distribution of the summary (scalar) scores would enter our model as μ_B.

A. Absolute Value of a Technology

Begin with the *ex ante* Taylor series expansion of the expected utility of the patient who receives the benefits of the new technology, expanding $u(M, H)$ around $u(M, \bar{H}_S + \mu_B)$ and scaling by EU_M. Thus:

$$E(V(H_S + \beta)) \approx p\left\{\frac{u(M,\bar{H}_S+\mu_B)}{EU_M} + \frac{1}{2}\frac{u_{HH}(M,\bar{H}_S+\mu_B)}{EU_M}\sigma^2_{S+B} + \frac{1}{6}\frac{u_{HHH}(M,\bar{H}_S+\mu_B)}{EU_M}\gamma_{1(S+B)}\sigma^3_{S+B} + \frac{1}{24}\frac{u_{HHHH}(M,\bar{H}_S+\mu_B)}{EU_M}\gamma_{2(S+B)}\sigma^4_{S+B} \ldots\right\} \quad (3)$$

The following Proposition, proven in the Appendix, provides sufficient conditions under which this Taylor Series converges.

Proposition 1: Suppose the health index, H, has support on the interval $[0, 1]$, and the utility function takes the form $U(c, H) = u(c)v(H)$, where v belongs to the class of HARA (hyperbolic absolute risk-aversion) utility functions. If H is a random variable with

support $[0, 1]$ and mean μ_0, then the Taylor expansion of $v(H)$ around μ_0 converges for all values in the support of H.

When health is represented as a quality-adjusted life-year, its domain is [0,1], as required by this Proposition. In addition, the HARA family of utility functions contains numerous commonly used forms: constant relative risk-aversion (power utility functions), constant absolute risk-aversion (exponential utility), increasing and decreasing absolute risk-aversion (IARA and DARA), increasing and decreasing relative risk aversion (IRRA and DRRA) and quadratic utility. HARA utility functions all exhibit the useful property that risk tolerance (the inverse of absolute risk-aversion) is linear in wealth [27].

The Taylor Series in Equation (3) allows us to express the utility function in terms of risk preference parameters known in the economics literature. Define the coefficient of absolute risk-aversion in consumption, $r \equiv -\frac{u_{MM}(M,\bar{H}_S)}{u_M(M,\bar{H}_S)}$. We always evaluate this, and the other risk preference parameters, at the baseline untreated health level, under the assumption that the health technology produces a marginal improvement in health that does not materially change the level of absolute risk-aversion. Define the coefficient of absolute prudence in consumption, $\pi \equiv -\frac{u_{MMM}(M,\bar{H}_S)}{u_{MM}(M,\bar{H}_S)}$, and absolute temperance as $\tau \equiv -\frac{u_{MMMM}(M,\bar{H}_S)}{u_{MMM}(M,\bar{H}_S)}$ [14, 28]. Analogously, define the coefficient of relative risk-aversion in consumption as $r^* \equiv rM$, relative prudence as $\pi^* \equiv \pi M$, and relative temperance as $\tau^* \equiv \tau M$. Box A discusses these terms in more detail.

[BOX A GOES HERE]

These terms characterize risk preferences with respect to consumption. In our application, we are more directly concerned with risk preferences over health. Define the coefficient of absolute risk-aversion in health as $r_H \equiv -\frac{u_{HH}(M,\bar{H}_S)}{u_H(M,\bar{H}_S)}$, of absolute prudence in health as $\pi_H \equiv$

$-\frac{u_{HHH}(M,\bar{H}_S)}{u_{HH}(M,\bar{H}_S)}$, and of absolute temperance in health as $\tau \equiv -\frac{u_{HHHH}(M,\bar{H}_S)}{u_{HHH}(M,\bar{H}_S)}$. Similarly, define relative risk aversion in health as $r_H^* \equiv r_H \bar{H}_S$, relative prudence in health as $\pi_H^* \equiv \pi_H \bar{H}_S$, and relative temperance in health as $\tau_H^* \equiv \tau_H \bar{H}_S$.

The economics literature provides estimates for relative risk preferences over consumption, but to date has not estimated relative risk preferences over health. Therefore, while our theoretical results are expressed in terms of the health risk preference parameters, our empirical applications impose the assumption that relative risk preferences over health are equal to relative risk preferences over consumption. This assumption is not required for our theoretical results, but it allows practitioners to incorporate risk preferences based on the current empirical literature, and it would be exactly correct in cases where $U(c,H) = v(c)v(H)$. Formally, we assume that $r^* = r_H^*$, $\pi^* = \pi_H^*$, and $\tau^* = \tau_H^*$.

Finally, observe that $\rho \equiv \frac{pu_M}{EU_M}$ is the marginal rate of substitution between a unit of consumption in the *ex ante* state and an expected unit in the *ex post* sick state [29]. This is equivalent to the willingness to pay ex ante for a $1 increase in the actuarial value of insurance payable in the sick state. With these parameters in hand, simple algebraic manipulation allows us to rewrite Equation (3) as:[5]

$$E(V(H_S+\beta)) \approx \rho\left\{\frac{u(M,\bar{H}_S+\mu_B)}{u_M} - \frac{1}{2}r_H K\sigma_{S+B}^2 + \frac{1}{6}\pi_H r_H K\,\gamma_{1(S+B)}\sigma_{S+B}^3 - \right.$$

$$\left.\frac{1}{24}\tau_H\pi_H r_H K\gamma_{2(S+B)}\sigma_{S+B}^4 \ldots\right\} \qquad (4)$$

The structure of this Taylor expansion has an intuitive interpretation: The scalar ρ converts consumption in the sick state into *ex ante* dollars. Thus, while $E(V(H_S+\beta))$ is

[5] Note that r, π, and τ each contain a negative sign. Thus the signs of the even power terms in (4), variance and kurtosis, are negative, but in the skewness term, the products of two negative terms create the positive sign.

expressed in terms of *ex ante* dollars, the term in curly braces is expressed in terms of sick state (*ex post*) consumption units.

The first term inside the curly braces, $\left(\frac{u(M,\bar{H}_S+\mu_B)}{u_M}\right)$, is the value of a statistical life-year spent at the average health level $\bar{H}_S + \mu_B$. This is the conventional estimate for the value of spending one year in the treated state, and it abstracts from uncertainty. The second term, $-\frac{1}{2}r_H K\sigma_{S+B}^2$, adjusts for the presence of variance in the health gain. Higher variance (σ_{S+B}^2) creates lower certainty-equivalent values of the treated state. Specifically, $\frac{1}{2}r_H K\sigma_{S+B}^2$ represents the risk premium associated with the use of the new technology, measured in units of consumption. The risk premium of any consumption gamble a is defined as $\frac{1}{2}rV(a)$, where r is the coefficient of absolute risk-aversion, and $V(a)$ is the variance of a [12, 13]. Thus, $\frac{1}{2}r_H\sigma_{S+T}^2$ measures the units of health the consumer would give up in exchange for eliminating the variance in the treated health distribution. Multiplying by K scales this risk premium in terms of consumption units.

Higher-order terms similarly account for the cost (or benefit) to consumers of changes in skewness and kurtosis that new technologies create. The term $\frac{1}{6}\pi_H r_H K\gamma_{1(S+B)}\sigma_{S+B}^3$ measures the consumption-denominated risk premium associated with $\gamma_{1(S+B)}$, skewness in the health outcome. The term $-\frac{1}{24}\tau_H\pi_H r_H K\gamma_{2(S+B)}\sigma_{S+B}^4$ measures the consumption-denominated risk premium associated with $\gamma_{2(S+B)}$ kurtosis in the distribution of health outcomes.

B. Incremental Value of a Technology

Equation (4) approximates the level of utility associated with the new technology, but we wish to characterize its incremental utility. Therefore, we perform a similar Taylor Series expansion of utility in the untreated sick state, $\frac{u(M,\bar{H}_S)}{EU_M}$, around $\bar{H}_S$:

$$E(V(H_S)) \approx \rho\left\{\frac{u(M,\bar{H}_S)}{u_M} - \frac{1}{2}r_H K\sigma_S^2 + \frac{1}{6}\pi_H r_H\, K\gamma_{1(S)}\sigma_S^3 - \frac{1}{24}\tau_H\pi_H r_H K\gamma_{2(S)}\sigma_S^4 \ldots\right\} \quad (5)$$

The difference between $E\big(V(H_S+\beta)\big)$ and $E(V(H_S))$ approximates a new technology's marginal value in units of *ex ante* consumption.[6] Maintaining our assumption of a marginal improvement in health, we treat absolute risk-preference parameters as roughly constant across treated and untreated states. This results in an approximation for $E(\Delta V(T))$, which we define as a technology's incremental value in terms of *ex ante* consumption:

$$E(\Delta V(T)) \approx \rho K\left\{\left(\frac{u(M,\bar{H}_S+\mu_B)-u(M,\bar{H}_S)}{u_M}\right)\frac{1}{K} - \frac{1}{2}r_H(\sigma_{S+B}^2-\sigma_S^2) + \frac{1}{6}\pi_H r_H\big(\gamma_{1(S+B)}\sigma_{S+B}^3 - \gamma_{1(S)}\sigma_S^3\big) - \frac{1}{24}\tau_H\pi_H r_H\big(\gamma_{2(S+B)}\sigma_{S+T}^4 - \gamma_{2(S)}\sigma_S^4\big)\ldots\right\} \quad (6)$$

To a first-order approximation, $\left(\frac{u(M,\bar{H}_S+\mu_B)-u(M,\bar{H}_S)}{u_M}\right)$ is the *ex post* willingness to pay for μ_B units of health; the approximation becomes exact for marginal improvements in health. Since this willingness to pay can be written equivalently as $K\mu_B$, we can eliminate the expressions for utility, using Δ to denote the changes in variance, skewness, and kurtosis between the treated and untreated states, resulting in:

$$E(\Delta V(T)) \approx \rho K\left\{\mu_B - \frac{1}{2}r_H\Delta\sigma^2 + \frac{1}{6}\pi_H r_H\Delta(\gamma_1\sigma^3) - \frac{1}{24}\tau_H\pi_H r_H\Delta(\gamma_2\sigma^4)\ldots\right\} \quad (7)$$

Note that in this formulation, $\Delta(\gamma_1\sigma^3)$ indicates the change in the third central moment, and similarly for the fourth-order term. As discussed above, the last three terms approximate the risk premium associated with this new technology, using not only changes in variance (associated with risk aversion), but also changes in skewness (associated with prudence), and changes in kurtosis (associated with temperance) induced by the new technology.

[6] Later, we study the convergence properties of this Taylor Series equations.

The expression in curly braces represents the certainty-equivalent gain in QALYs experienced by users of the technology. If the actual QALY gain is stochastic, this term yields the non-stochastic gain in QALYs that would be equivalent to it. For example, a stochastic gain of 2.0 average QALYs might be worth 1.5 non-stochastic or "certainty-equivalent" QALYs, if it involves considerable variance.

This certainty-equivalent QALY gain consists of components that depend on mean and higher-order moments of the QALY gain. The first term—μ_B—gives the standard mean-based measure of QALYs gained. On its own, the second term encapsulates the gain (reduction) in equivalent units of health that the risk-averse consumer enjoys (suffers) if the technology decreases (increases) the variance of health outcomes. Earlier research has defined this as the "insurance value" of medical technology [6]. New technologies that lower the overall variance of sick state health outcome ($H_s + \beta$) generate the most insurance value to consumers. For instance, technologies that effectively treat the most severe illness state – i.e., those possessing negative covariance $\sigma_{S\beta}$ – are most valuable, because they reduce the variance of $H_s + \beta$ the most. In addition, technologies with the least variable treatment effects— i.e., lower values of σ^2_{S+B}— are the most-valuable to a risk-averse consumer. Thus, treatments that target severe states and that produce consistent health improvements produce the most insurance value to risk-averse consumers.

The third term implies that positively skewed treatment outcomes are potentially valuable to risk-averse consumers. This result sheds light on prior evidence suggesting that patients might prefer treatments with more positively skewed treatment effects. Several studies have presented patients with a stated-preference choice between two therapies that have the same average treatment benefit, where one treatment has a non-random outcome and the other has the same

expected health gain but more variance and positive skew [30, 31]. In these settings, patients have expressed a preference for the second type of therapy. This is consistent with patients responding to increases in positive skew rather than (and perhaps even in spite of) increases in variance. Prior studies have inferred from this phenomenon the "value of hope," a preference for modest chances at a positively skewed outcome. Holding variance constant, as skewness changes, $\Delta\gamma_{1(S+B)}$ increases, and so does this "value of hope" for consumers. Note that in our formulation, the Pearson skewness parameter is multiplied by σ^3, reflecting the interaction of skewness and variance.

The fourth order term involving kurtosis may be more problematic to incorporate into real-world estimates of value. Obtaining sufficiently precise estimates of kurtosis may expand sample sizes in clinical trials to unreasonable levels, particularly because we require differences in kurtosis, the variance of which is the sum of the individual variances of the separate estimates of kurtosis. If available data exist, analysts can readily incorporate these effects into their analysis. Pearson's kurtosis measure is multiplied by σ^4 in our expressions involving the fourth central moment of the outcome distributions.

Empirical applications commonly employ the assumption that relative risk-aversion, relative prudence, and relative temperance are approximately independent of the level of consumption [32]. Therefore, it is convenient to reformulate Equation (7) in terms of relative rather than absolute risk preferences. This transformation has the added benefit of removing the dimensionality from the risk preference terms, which will no longer be sensitive to the magnitudes of health. Along the same lines, it is also useful to normalize the statistical central moments so that they do not vary with the scale of the dispersion in treatment outcomes. This allows us to perform the following transformation on Equation (7), recalling that $\bar{H}_S \equiv H_W - \mu_S$:

$$(\Delta V(T)) \approx \rho K \mu_B \left[1 - \frac{1}{2} r_H^* \left(\frac{\mu_B}{\bar{H}_S}\right)\left(\frac{\Delta \sigma^2}{\mu_B^2}\right) + \frac{1}{6}\pi_H^* r_H^* \left(\frac{\mu_B^2}{\bar{H}_S^2}\right)\left(\frac{\Delta(\gamma_1 \sigma^3)}{\mu_B^3}\right) - \right.$$

$$\left. \frac{1}{24}\tau_H^* \pi_H^* r_H^* \left(\frac{\mu_B^3}{\bar{H}_S^3}\right)\left(\frac{\Delta(\gamma_2 \sigma^4)}{\mu_B^4}\right) + \cdots\right] \qquad (8)$$

This represents the *ex ante* value in consumption units of the QALY gain from the technology. The term in square braces in Equation (8) represents the "certainty-equivalence ratio," which adjusts the average QALY gain to account for differences in risk profiles of the new and comparison technologies. We define this ratio as ϵ_Δ, which indexes the proportional amount of compensation that risk-averse individuals would demand in exchange for bearing the incremental risk associated with the new medical technology. For example, if $\epsilon_\Delta = 0.8$, individuals would demand a payment (or reduction in price) equal to 20% of the value of the mean improvement in health. In parallel, if $\epsilon_\Delta = 1.2$, a risk averse person would be willing to pay 20% more than for an otherwise-similar technology that created no risk reduction.

The ratio ϵ_Δ is defined by:

$$\epsilon_\Delta \approx \left[1 - \frac{1}{2} r_H^* \left(\frac{\mu_B}{\bar{H}_S}\right)\left(\frac{\Delta\sigma^2}{\mu_B^2}\right) + \frac{1}{6}\pi_H^* r_H^* \left(\frac{\mu_B^2}{\bar{H}_S^2}\right)\left(\frac{\Delta(\gamma_1 \sigma^3)}{\mu_B^3}\right) - \frac{1}{24}\tau_H^* \pi_H^* r_H^* \left(\frac{\mu_B^3}{\bar{H}_S^3}\right)\left(\frac{\Delta(\gamma_2 \sigma^4)}{\mu_B^4}\right) + \cdots\right] \quad (9)$$

The product $\epsilon_\Delta \mu_B$ represents the certain gain in QALYs that is equal in value to the risky gain offered by the technology. By definition, $\epsilon_\Delta \approx 1$ if treatment outcomes are nearly certain in both the treated and untreated states, if they do not differ meaningfully in their stochastic components, or if consumers are risk-neutral. The extent to which ϵ_Δ differs from one (in either direction) indicates the percent error when the analysis ignores the stochastic components of the proper expression for expected utility.

Equation (9) captures three terms expressing the effect of uncertainty on economic value (relative to the customary mean-based measure of value). The first term (involving relative risk aversion in health, r_H^*, and the change in variance of outcomes) is comparable to the standard

Pratt-Arrow measure of risk aversion. The second term, involving r_H^*, relative prudence, π_H^*, and differences in the third central moment values expresses the "value of hope" that right-skewed distributions of outcomes can produce. The third term, involving r_H^*, π_H^*, and τ_H^* (relative temperance) and the change in kurtosis, moderates the effects of changes in variance and skewness. Generally, if variance and kurtosis change in the same direction, this amplifies the effects of variance changes, and conversely. Similarly, when changes in skewness and kurtosis move in opposite directions, that amplifies the effect of skewness changes, and conversely.

We can reformulate Equation (9) to demonstrate two key issues associated with our model:

$$\epsilon_\Delta \approx \left[1 - \{\frac{1}{2} r_H^* \left(\frac{1}{\overline{H}_S}\right) \Delta\sigma^2 + \frac{1}{6} \pi_H^* r_H^* \left(\frac{1}{\overline{H}_S^2}\right) \gamma_1 \Delta\sigma^3 - \frac{1}{24} \tau_H^* \pi_H^* r_H^* \left(\frac{1}{\overline{H}_S^3}\right) \gamma_2 \Delta\sigma^4 + \cdots\} (\frac{1}{\mu_B})\right]$$

$$\equiv 1 + \frac{\{Risk\ Components\}}{\mu_B} \qquad (10)$$

First, note that the relative importance of the risk components diminishes as μ_B increases. Intuitively, this makes good sense. As the mean difference increases, the stochastic (higher order) terms in the Taylor Series have less proportional consequence. Conversely, this also implies that when mean differences in treatment outcomes are relatively small, the stochastic information becomes more important. We would expect that over time, incremental gains in treatment efficacy for a given disease condition will become increasingly more difficult to attain, hence the importance of the stochastic components may well rise over time [33]. Subsequent simulations will demonstrate this phenomenon.

The second observation arises from our assumption of constant relative risk aversion parameters (r_H^*, π_H^* and τ_H^*). This formulation adds increasingly higher powers of $\overline{H}_S$ in the denominators of the risk components (in order to create relative risk aversion parameters defined

in terms of health in the untreated sick state). This makes the risk-related terms relatively more important for more severe diseases, with values of $\overline{H}_S$ closer to zero. We illustrate the quantitative implications in our simulations. As we demonstrate in additional simulations, this also means that convergence of the Taylor Series hinges to some extent on the assumed values of $\overline{H}_S$. In general, $0 < \overline{H}_S < 1$, with 0 representing death and 1 representing perfect health. However, even a terminal illness diagnosis with just a few months of remaining life could give $\overline{H}_S$ values of 0.2 to 0.5, for example, so extremely small values of $\overline{H}_S$ are not particularly relevant to our analysis.

IV. IMPLICATIONS FOR COST-EFFECTIVENESS ANALYSIS

We now turn to the performance of cost-effectiveness analysis for medical technologies of uncertain benefit. Without treatment, patients experience the random health outcome, $H_S \equiv H_W - s$. With the new technology, they experience the random outcome, $H_S + \beta$. Recall that μ_B is the mean health benefit that results from the use of this technology. This parameter would be estimated by computing the difference in mean outcomes across the treatment and control groups in a randomized trial of the medical technology. Conventionally, cost-effectiveness analysis calculates this mean improvement due to a healthcare technology and compares the incremental cost-effectiveness ratio (ICER) with the operational willingness to pay for QALYs (K). An alternative approach selects a specific value of K and monetizes the gain—the "net monetary benefit" (NMB) form of a cost-effectiveness calculation. In our terminology, the standard framework produces the net monetary benefit from the *ex ante* perspective as simply $K\rho\mu_B$. Our analysis generalizes this conventional approach and incorporates stochastic health improvement by including higher-order terms in the Taylor Series expansion of expected utility.

Consider the technology adoption decision. Conceptually, the question is whether it is efficient for a health insurer to cover the technology. This question is best viewed from the *ex ante* perspective of an insured consumer whose health status is not yet realized. Define $\Delta C(T)$ as the incremental cost of technology T when administered to a sick patient. The actuarially fair *ex ante* increase in insurance premium costs of covering the technology is $p\Delta C(T)$. Therefore, the technology is adopted if $\Delta V(T) \geq p\Delta C(T)$, or equivalently, if $K\rho\mu_B\epsilon_Q(T) \geq p\Delta C(T)$. The latter condition is equivalent to:

$$\frac{p\Delta C(z)}{\rho\mu_B\epsilon_\Delta(T)} \leq K \tag{11}$$

The left-hand side of this inequality is the ratio of ex ante incremental costs to ex ante certainty-equivalent QALYs gained. This is a more general version of the typical incremental cost-effectiveness ratio (ICER), adjusted here to account for stochastic changes in QALYs from the *ex ante* perspective of an individual whose health status has not yet been realized. Our framework nests the traditional ICER decision rule as the special case in which $\epsilon_\Delta(T) = 1$. Put differently, by ignoring the stochastic elements of the analysis, traditional cost-effectiveness analysis presumes that $\epsilon_\Delta(T) \equiv 1$.

It is important to note that incorporating the stochastic nature of treatment effects does not require development of a new maximum WTP cutoff value K. Derivation of the optimal level of K used a framework of expected utility maximization [10], and our Taylor Series expansion provides a more accurate representation of changes in expected utility than simply using the average improvement in quality of life outcomes (μ_B). Estimates of the optimal value of K thus apply directly to our approach to estimating the value of health technologies with stochastic outcomes [21].

Several extensions to our model are relatively straightforward. In the first place, our model allows for incorporating adverse events from medical treatments (AEMTs) in two ways. First, health and cost consequences of AEMTs could be summed directly into net health gains and net cost changes. Doing so would imply that the health consequences of AEMTs should be valued identically to other health changes. An argument exists that AEMTs should be valued at willingness to accept (WTA) rather than willingness to pay (WTP) valuations, where WTA > WTP, perhaps by considerable amounts [34] . Our model would readily incorporate such concepts by separately accumulating AEMT risks, and valuing them at νK, where $\nu > 1$.

In addition, the analysis above presumes technology adoption decisions by a social planner. In contrast, a payer with a fixed healthcare budget that falls below the efficient level will solve a constrained maximization problem with a binding budget constraint. For such a payer, the threshold, K, will be more stringent but the rest of the problem is formulated identically.

V. INCORPORATING LONGEVITY EFFECTS

The analysis so far has focused on changes in quality of life. We now incorporate changes in probabilities of survival as well.

A. Life-Cycle Framework

Suppose that our health technology, T, produces stochastic survival benefits, along with its stochastic quality of life benefits. The probability of a treated patient surviving from period zero to period i is given by the random variable, λ_{Ti}, which has expected value $E(\lambda_{Ti})$, essentially equivalent to actuarial life table values. The treated patient's period i utility is given by $u(M_i, H_{Si} + \beta_i)$. Note that mortality risk does not influence the period utility function. To abstract from the dynamics of intertemporal consumption smoothing, there is no borrowing or

lending in our model, so mortality has no effect on the level of consumption.[7] Define $\delta(\lambda_{Ti})$ as the density of λ_{Ti}. To conserve notation, define $Y_{Ti} \equiv (M_i, H_{Si} + \beta_i)$. The treated individual with the one-period utility discount factor ϕ derives expected lifetime utility:

$$u(Y_{T0}) + \sum_{i=1}^{\infty} \left(\frac{1}{1+\phi}\right)^i \int_{\lambda_{Ti} \in \Lambda_i} \lambda_{Ti} u(Y_{Ti}) \delta(\lambda_{Ti}) d\lambda_{Ti} = u(Y_{T0}) + \sum_{i=1}^{\infty} \left(\frac{1}{1+\phi}\right)^i E(\lambda_{Ti}) u(Y_{Ti}) \quad (12)$$

The survival probability itself is allowed to be stochastic in this framework. In the conventional case where survival probabilities are modeled deterministically, $\delta(\lambda_{Ti}) \equiv 1$.

Observe that we maintain the typical life-cycle framework in which period utilities are additively separable. Thus, since expected utility is linear in mortality risk, individuals exhibit risk-neutrality over mortality risk.[8] This contrasts with consumer risk-aversion over quality of life. It also eliminates one strong implication of the common quality-adjusted life-years formulation, namely that a one percent decrease in mortality risk is just as valuable as a one percent increase in quality of life. This conventional implication would persist only if risk-aversion over quality of life is identical to risk-aversion over mortality.

B. Absolute Value of Technology

As a result of risk-neutrality over mortality, only expected mortality risk enters the consumer's lifetime utility-maximization problem. Define λ_{Si} as the stochastic survival rate in period i under the standard of care, and define the expected survival benefit of the treatment as

[7] The conventional life-cycle model for the value of a statistical life also implies independence between mortality risk and life-cycle consumption decisions, although in this case the result arises due to perfect annuitization [1, 2]. The presence of incomplete financial markets can produce a relationship between mortality risk and consumption [35].

[8] Some authors have explored various alternative formulations that introduce risk-aversion over longevity [cf, 36, 37]. We remain agnostic about the correct formulation and proceed using the standard model in order to focus attention on other issues. Regardless, even under the alternative formulations, there is no reason to believe that the degree of risk-aversion over quality of life is necessarily identical to that over quantity of life. As such, our conclusions regarding the challenges of the QALY formulation remain.

$\Delta E(\lambda_i) = E(\lambda_{Ti}) - E(\lambda_{Si})$. The treatment now has a two-dimensional benefit in each period i, $(\Delta E(\lambda_i), \mu_{Bi})$, and its value can be written as the sum of two parts: (1) the change in survival probability multiplied by the value of a period i life-year spent in the baseline untreated state; plus (2) the expected probability of surviving to period i, multiplied by the certainty-equivalent value of morbidity reduction in period i. This results in:

$$E(\Delta V_i(\Delta E(\lambda_i), \mu_{Bi}, H_{Si})) \equiv \Delta E(\lambda_i) V(H_{Si}) + E(\lambda_{Si} \Delta V(T)) \qquad (13)$$

Relying on the earlier development of the term $E(\Delta V(T))$, we can rewrite the second term in this expression as $E(\lambda_{Si})\rho K \epsilon_\Delta(T)\mu_B$. Since $\epsilon_\Delta(T)\mu_B$ is the certainty-equivalent gain in QALYs associated with the new technology, the product $\rho K \epsilon_\Delta(T)\mu_B$ is the value (in ex ante consumption units) of that certainty-equivalent gain. Finally, $E(\lambda_{Si})\rho K \epsilon_\Delta(T)\mu_B$ is the expected value of the certainty-equivalent gain in QALYs, discounted by the risk of dying before the QALY gain is experienced.

Employing this logic, we can rewrite equation (13) so that it measures the expected value of certainty-equivalent QALYs gained.

$$E(V(H_{si})) \approx \rho K \left\{ \left[\frac{u(M_i, \bar{H}_{Si})}{u_M K \bar{H}_{si}} - \frac{1}{2} r_H^* \frac{\sigma_{Si}^2}{\bar{H}_{Si}^2} + \frac{1}{6} \pi_H^* r_H^* \frac{\gamma_{1si} \sigma_{Si}^3}{\bar{H}_{Si}^3} - \frac{1}{24} \tau_H^* \pi_H^* r_H^* \frac{\gamma_{2Si} \sigma_{Si}^4}{\bar{H}_{Si}^4} \ldots \right] \bar{H}_{si} \right\} \quad (14)$$

Note that, comparing this expression to $v(H_{si} + \tau)$, the terms $\frac{\mu_B}{\bar{H}_S}$ drop out, because we are normalizing absolute risk preferences and the statistical moments by the same parameter, $\bar{H}_{Si}$. Notice also that the term in square brackets equals $\frac{V(H_{Si})}{\rho K \bar{H}_{si}}$, the certainty-equivalence ratio in the level of health, as opposed to the gain in health. We define it as $\epsilon(H_{Si})$.

To summarize this concept, every one-unit gain in average QALYs is worth $\epsilon(H_{Si})$ QALYs to the consumer. When health is stochastic and consumers are risk-averse over health, they demand compensation for the variance in H_{Si}. As a result, $\epsilon(H_{Si}) < 1$. On the other hand,

positive skewness in H_{Si} could result in $\epsilon(H_{Si}) > 1$, so that average QALY gains might understate the value of health improvements.

With the definition of $\epsilon(H_{Si})$ in hand, we can rewrite the value of the technology for period i as:

$$E(\Delta V_i(\lambda_i, \mu_{Ti}, H_{Si})) \equiv \rho K\{\Delta E(\lambda_i)\epsilon(H_{Si})\bar{H}_{Si} + E(\lambda_{Si})\epsilon_{\Delta}(T)\mu_{Bi}\} \quad (15)$$

We define the term in curly braces as the quality- and risk-adjusted life-years (QRALYs) gained in period i as a result of the technology:

$$\Delta QRALY_{\mathrm{i}} = \Delta E(\lambda_i)\epsilon(H_{Si})\bar{H}_{Si} + E(\lambda_{Si})\epsilon_{\Delta}(T)\mu_{Bi} \quad (16)$$

$QRALYs$ gained equal the certainty-equivalent QALY gain, discounted by the probability of dying before the gain is accrued.

Conveniently, our expressions nest traditional cost-effectiveness as a special case. Observe that in the traditional model, average QALYs gained equal certainty-equivalent QALYs gained, and $\epsilon(H_{Si}) = 1$. The traditional model assumes that health is non-stochastic and that $u(M, H) \equiv v(M)H$. These points imply that $\epsilon(H_{Si}) = \frac{u(M_i, H_{Si})}{u_M K H_{Si}}$ and $K = \frac{u_H}{u_M} = \frac{v(M)}{v'(M)H}$.

Therefore, $\epsilon(H_{Si}) = \frac{v(M)H_{Si}}{(v'(M)H_{Si})\left(\frac{v(M)}{v'(M)H_{Si}}\right)H_{Si}} = 1$ in the traditional cost-effectiveness model.

Moreover, recall that $\epsilon_{\Delta}(T) = 1$ in the traditional model, because average QALY gains are equal to certainty-equivalent QALY gains. Since $\epsilon(H_{Si}) = \epsilon_{\Delta}(T) = 1$, , it follows that gains in QRALYs are identical to gains in QALYs. As a result, all our results prove identical to the standard cost-effectiveness framework in the special case where QRALYs and QALYs are equal.

Returning to the generalized QRALYs framework, the net monetary benefit of the technology in period i is:

$$NMB_i = \rho K \Delta QRALY_i \quad (17)$$

Extending to the multi-period case, assume that units of consumption are discounted at the rate of interest $-\phi$. Technology adoption is welfare-improving whenever its lifetime net present value to the consumer is nonnegative, or its total discounted net monetary benefit is positive:[9]

$$PDV(\{NMB_i\}_{i=1}^{\infty}) \equiv \sum_{i=1}^{\infty} \left(\frac{1}{1+\phi}\right)^i [\rho K \, \Delta \, QRALY_i - C_i] \geq 0 \tag{18}$$

C. Incremental Cost-Effectiveness Ratios

We now develop incremental cost-effectiveness ratios of the usual form. If the technology costs C_i in each period i, it is (weakly) welfare-improving in period i if $\Delta V_i(\Delta E(\lambda_i), \mu_{Bi}, H_{Si}) \geq C_i$, or:

$$\frac{C_i}{\rho \Delta QRALY_i} \leq K \tag{19}$$

The generalized incremental cost-effectiveness ratio, on the left-hand side, calculates cost per QRALY gained, from the *ex ante* perspective.

There is an analogous multi-period decision rule. Making the typical assumption that the $ex\ ante$ value of QALYs is constant over time, the technology is welfare-improving over a lifetime horizon if $K \sum_{i=1}^{\infty} \left(\frac{1}{1+\phi}\right)^i \rho \Delta QRALY_i \geq \sum_{i=1}^{\infty} \left(\frac{1}{1+\phi}\right)^i C_i$. Therefore, the technology is welfare-improving if and only if:

$$\frac{\sum_{i=1}^{\infty} \left(\frac{1}{1+\phi}\right)^i C_i}{\sum_{i=1}^{\infty} \left(\frac{1}{1+\phi}\right)^i \rho \Delta QRALY_i} \leq K \tag{20}$$

This is again analogous to the typical multi-period decision rule, but replacing QALYs gained with QRALYs gained.

[9] To conserve notation, we assume the consumer discounts utility and consumption at the same rate, but this assumption is easily generalized.

VI. NUMERICAL SIMULATION METHODS AND PARAMETER ESTIMATES

In this section we present our methods for simulating how ϵ_Δ – our certainty-equivalence ratio (CER) — changes as different parameters of the distributions of health outcomes change. These simulations focus on changes in quality of life in a single period to compactly demonstrate our methodology. Similar results would appear in the multi-period formulations of Section V, but with the added complications of multiple years and possibly time-varying stochastic parameters.

The term ϵ_Δ neatly summarizes the implications of our theory, because it provides the ratio between our value estimate and that of the conventional, non-stochastic approach. We use Equation (9) as the basis for our simulations, graphing ϵ_Δ against changes in skewness for various levels of changes in variance, ignoring higher-order terms for visual and expositional clarity. In our initial simulations, we independently vary the changes in variance and skewness. Including effects of changes in kurtosis would complicate the analysis with little gain in intuition for the reader. In general, if kurtosis and variance move in the same direction, the omitted terms would amplify the effects of changing variance, and conversely. Similarly, if skewness and kurtosis move in opposite directions, that amplifies the effect of skewness changes, and conversely. Later, we also present simulations that calculate the speed at which the Taylor Series converges, using variance, skew, and kurtosis.

Our simulations independently vary the changes in standardized variance and standardized skewness to demonstrate how these measures affect the CER. Since we do not know of any reported measures of skewness in clinical trials, we cannot know for certain the potential range of standardized variances and skewness parameters (or more importantly, their

differences). The Figures that follow limit the range of changes in standardized variance and skewness in arbitrary ways for visual clarity. They also provide guidance to practitioners regarding how the CER varies with the moments of treatment effects, and how quickly the relevant Taylor Series converges.

To simulate the effects of higher-order terms in the Taylor Series expansion, we require estimates of the relative risk preference parameters, r^*, π^*, and τ^*. We also require estimates of the optimal willingness to pay (K) and the marginal rate of substitution between health and sick states (ρ) for the full measures of value. We discuss each of these in turn.

A. Estimates of Relative Risk-Aversion

Relative risk aversion has long been studied in the economic literature. Phelps (2019) summarizes central estimates for r^* as follows (see Section III for details and references) [21]:

Labor Supply-based estimates	$0.7 \leq r^* \leq 1$
Micro-simulation parameters	"about 0.8"
Studies of reported happiness	0.92 to 1.26 in developed nations
Experimental data	0.88 to 1.43 (centering on 0.93).

Taken together, these suggest that r^* is slightly below 1. We use $r^* = 1$ in our simulations.

B. Estimates of Relative Prudence and Temperance

For relative prudence and temperance, Noussair et al. (Table 11) provide a range of estimates with differently scaled lotteries and assumptions about the specific utility function [32]. Summarizing these parameters, we use $\pi^* = 1.75$ and $\tau^* = 2.5$ in our simulations.

Subsequent research estimates that π^* is close to 2.0 [38], although our results are not heavily sensitive to these parameter choices.[10]

We can also calibrate the estimate of relative prudence (π^*) from observations about the nature of risk aversion from financial portfolio selection [39]. Arrow observed that the *magnitude* of risky assets rises with wealth, so utility must exhibit decreasing absolute risk aversion (DARA). Separately, he observed that the *proportion* of portfolios devoted to risky assets declines as wealth rises. Now define the income elasticity of r as α. DARA requires $\alpha < 0$. IRRA requires $\alpha > -1$. Taken together, these require $0 > \alpha > -1$. This in turn requires $1 > (\pi^* - r^*) > 0$. The estimates of Noussair et al. (their Table 11) meet all these requirements [32].[11]

C. Optimal WTP Relative to Income and ρ

For formulae involving total value of a technology (but not ϵ_Δ), we also need empirical values for K (the optimal WTP for health gains) and ρ (the *ex ante* to *ex post* conversion factor) to estimate the incremental value of a new technology (but not to estimate ϵ_Δ). A variety of ways exist to estimate the optimal K, the most common of which relies on wage premiums for risky occupations. A new method for estimating $\frac{K}{M}$ uses a specific utility function (the Weibull function, also known as expo-power utility), calibrated using estimates of relative risk aversion (r^*). That approach yields an estimate of $\frac{K}{M}$ of approximately 2 at average levels of income in the US and other developed countries, and also shows that the optimal levels of both K and $\frac{K}{M}$ increase with income [21]. In assessing total value equations, we would use $\frac{K}{M}$ = 2 as necessary.

---

[10] For comparison, in the power utility function, r* = 1, $\pi^* = 2, \tau^* = 3, \ldots$.

[11] The proof begins by defining the elasticity of r with respect to M as α. Next, DARA requires $\alpha < 0$. Then it is easy to show that $(r^* - \alpha) = \pi^*$ Thus it must be true that $1 > (\pi^* - r^*) > 0$ if utility is both DARA and IRRA.

Recall that $\rho \equiv \frac{pu_M}{EU_M}$, the ex ante willingness to pay for a \$1 increase in the actuarial value of insurance payable in the sick state. This parameter has been estimated in the context of health insurance using discrete choice experiment survey methods [40]. In particular, a representative sample of American consumers was surveyed to determine how much additional insurance premium each would pay in exchange for \$1 of additional actuarial value payable in the sick state from a health insurance policy. The estimated willingness to pay was approximately \$2.50, which we would use in any equation where ρ appears.

VII. SIMULATION RESULTS

A. Independent Changes in Variance and Skewness

We first graph ϵ_Δ in the stylized context where variance and skewness vary independently. While standard statistical distributions do not exhibit this independence between variance and skew, this exercise facilitates intuition and insight. The basic structure of Figure 1 follows Equation (9), truncated after the third term (skewness). The vertical axis is the CER ratio ϵ_Δ, graphed against the change in standardized skewness for various changes in standardized variance. Table 1 summarizes the general effect of changes in the key parameters – variance and skewness of T and S.

Table 1 here

In Figure 1, the topmost lines represent larger reductions in variance of T compared with S, so that the value of $\frac{\Delta\sigma^2}{\mu_B^2}$is negative (more so for higher lines), and for the lower lines, the variance has increased. Decreased (increased) variance is an economic good (bad). The heavy horizontal line at $\epsilon_\Delta = 1$ is the "breakeven" for the stochastic terms. Figure 1 contains four panels with increasing values of $\mu_B = \{0.1, 0.2, 0.3, 0.5\}$ and a fixed value of $\overline{H}_s = 0.5$.

[FIGURES 1a – 1d here]

Several things are apparent from these graphs. First, as the health gain μ_B increases, the relative importance of the stochastic terms decreases (as shown in Equation 10) and the group of lines flattens out. We should expect nothing else. At the extreme, if two technologies had identical means ($\mu_B = 0$), their stochastic profiles alone would determine which was preferable (see Equation (7)). In panel (a) of Figure 1, with $\mu_B = 0.1$, the highest values of ϵ_Δ are near 1.6 for a change in standardized Pearson skewness of 0.5. For the largest average health gain in Panel (c), ($\mu_B = 0.3$) ϵ_Δ values are below 1.3 for otherwise-identical parameterization of the model. When μ_B reaches 0.5 (Panel (d)), the upper line (the largest difference in variances) is essentially flat.

Second, within any panel in Figure 1, the lines have different slopes because the horizontal axis graphs the Pearson skewness parameter (a commonly reported statistical parameter) rather than the third central moment that our model requires. This introduces an interaction between the slopes of the lines and the magnitude of σ_{B+S}. As the variance of T falls, the difference in variances rises (holding σ_s^2 constant), so the upper lines in all panels of Figure 1 have flatter slopes than the lower lines in each panel. Asymptotically, as σ_{B+S} approaches 0, for any given value of σ_S, the slope of these lines also approaches 0.

[FIGURES 2a –2d here]

Figure 2 performs a similar exercise as Figure 1, but for variations in baseline health status, $\bar{H}_S$. This figure fixes μ_B at 0.25 and varies $\bar{H}_S$ from 0.2 to 0.8 for the same range of variation in skewness and variance shown in Figure 1. Lower levels of baseline health ($\bar{H}_S$) magnify the effects of variance and skewness. For baseline health, $\bar{H}_S = 0.2$, the term ϵ_Δ ranges from about -2.0 to 4.0. In contrast, when $\bar{H}_S = 0.8$, ϵ_Δ ranges from around 0.75 to 1.2.

Intuitively, risk-averse people find it costlier to bear risk when their marginal utility is higher; this explains why there is a greater demand for certainty in worse health states.

For highly severe illness, neglecting uncertainty can misstate value by a factor of two to four. In cases where treatment effects are negatively skewed, this could even cause the sign of the estimate to be wrong. To be sure, the latter case is fairly extreme, as $\bar{H}_S = 0.2$ is an extremely severe illness, close to death in its seriousness. Nevertheless, this analysis suggests that conventional CEA understates the relative value of effectively treating severe illness compared to milder disease. The degree of understatement rises with the improvement in variance and skewness in the treatment effect.

These simulations make several things evident. Most importantly, ϵ_Δ (the CER) can vary substantially from a value of 1 when stochastic components of expected utility are taken into account, and especially for severe illness. Shifts away from $\epsilon_\Delta = 1$ occur both from changes in variance and skewness. Omitting these considerations can significantly bias technology evaluations in either direction, depending on how the new technology changes the risk profile for treated patients. Further, the bias is more important with relatively small changes in the mean health benefit.

B. Examples with 4-parameter Beta distributions.

To provide additional examples of our model, we use four-parameter Beta distributions to describe distributions of health outcomes for treated and untreated patients. Box B discusses this distribution for those not familiar with it. Beyond the standard two-parameter Beta distribution, two additional parameters (c and d) define the region of support, which we here limit to subsets of the $[0,1]$ interval. Our simulations in Figures 3 and 4 focus on distributions where T has support in the $[0.5,1]$ range (width of support (c-d) = 0.5) and S has support with a width of (c-d)

= 0.4. The other parameters (α and β) define the mean, variance, skewness and kurtosis, as defined in Box B. Specific parameters appear in the legends of each figure.

BOX B here

Figures 3a-d show the effect of reducing variance and increasing skewness of the distribution of the new therapy (T) relative to the comparison therapy, which always has a mean outcome of 0.6 QALYs, a fixed variance (0.011) and zero skewness. Technology T has a ratio of α and β fixed at 1:3, thereby introducing moderate skewness. As the absolute value of those parameters falls, the mean remains unchanged but the variance and skewness both increase. Higher variance reduces the stochastic terms, because risk-averse consumers dislike it. This is reflected in a lower certainty-equivalence ratio, ϵ_Δ, because more variable treatments result in lower certainty-equivalent QALY gains. Higher positive skew does the opposite, because prudent consumers value positive skew. Unfortunately, we cannot independently alter the skewness and variance in the 4-parameter Beta distribution, nor in any other probability distribution of which we are aware. But since they change at different rates, we can nonetheless observe the separate effects of variance and skewness.

Figure 3 here

Figure 3a, ($\alpha = 4$ and $\beta = 12$) has the smallest variance for T and the largest certainty-equivalence ratio, with $\epsilon_\Delta = 1.46$. In this case, nearly 99% of the stochastic effect comes from variance rather than skew. Reducing the parameters (Figure 3b) to $\alpha = 2$ and $\beta = 6$ increases the variance and skewness such that $\epsilon_\Delta = 1.34$. Compared with Figure 3a, the contribution of skew to the stochastic terms rises from 1.2% to 5.5%. Reducing these parameters to $\alpha = 1.2$ and $\beta = 3.6$, yields $\epsilon_\Delta = 1.23$. As the variance of T rises, the certainty-equivalence ratio falls; the certainty-equivalent gain in QALYs is smaller for a more variable treatment. Dropping

the parameters to $\alpha = 0.7$ and $\beta = 2.1$ increases the variance of T to slightly above that of S. This would otherwise lead to a value of $\epsilon_\Delta < 1$, except for the positive skewness, which gives a net value of $\epsilon_\Delta = 1.07$, and 144% of the value of the stochastic terms comes from the positive skewness.

Figure 4 here

Figure 4 demonstrates the implications of higher mean treatment effects. Figure 4a reproduces Figure 3a, with a value of $\epsilon_\Delta = 1.46$. We then successively shift the distribution of the control therapy (S) leftward to increase the mean difference between T and S, holding all other parameters constant. The value of ϵ_Δ falls from 1.46 down to 1.07 as the mean treatment effect expands from $\mu_B = 0.025$ to $\mu_B = 0.175$. These simulations provide a separate visualization of the phenomenon discussed previously—the importance of the stochastic components of value is greater when only small incremental changes in average benefit occur between competing therapies.

One conclusion – or at least suggestion – from these simulations is that variance accounts for the lion's share of the value created by stochastic components in many situations. Particularly when ϵ_Δ is large (e.g., over 1.3 or more), improved variance accounts for a large fraction of the overall stochastic gain, as appears most strongly in Figures 3a and 3b. As the variance of T rises (and thus ϵ_Δ falls towards 1.0), the contribution of skewness increases in percentage terms, and can actually reach the situation where positive skewness overcomes what would otherwise be a loss in value from increased variance in T.

This suggests that as least in some cases, knowing variances of the distributions of treatment outcomes may present a reasonably clear picture of the overall stochastic gain (or loss) arising from the stochastic terms in our model. Thus, measuring skewness may not be highly

important in some settings. If true, this is in some sense good news, since randomized controlled trials already contain estimates of variances in different treatment arms, although standard reporting may obscure such data by only reporting p-values for differences of means. However, only experience with real-world data will reveal the extent to which this is a general phenomenon or merely an artifact of our simulations.

This result would not be overly surprising given the general nature of Taylor Series expansions, where the first and second-order terms often capture much of the overall estimate. In such cases, we also know the direction of bias, namely that omitting measures of skewness from Taylor Series estimates will understate or overstate the value of a new technology as the omitted change in skewness is positive or negative.

C. Assessing Convergence of the Taylor Series Expansion

Next, we assess the convergence characteristics of our Taylor Series expansion using four-parameter Beta distributions with varying skewness and kurtosis. We calculated the absolute value of the second, third and fourth-order Taylor series terms to compare various technologies (T) with a common control (S), where the control intervention always has a four parameter Beta distribution with parameters $\alpha = 2, \beta = 2$ (the zero-skewness distribution used in Figures 3 and 4) and support that is 0.5 units wide. Similarly, the T distributions feature $\alpha = 2$ and vary β between 5 and 10 in increments of 1, with 0.5 unit width of support. Each increase in β increases the skewness and alters the kurtosis in complex ways (since β appears in both the numerator and denominator of the expression for kurtosis).

We characterize speed of convergence by calculating ratios between the 3rd-order and 2nd-order Taylor Series terms, and between 4th-order and 2nd-order Taylor Series terms. Our simulations illustrate the result from Equation (10), i.e., that the speed of convergence varies

with the average level of health in the untreated sick state ($\overline{H}_S$). Figure 5a shows the ratio of the 3rd to the 2nd order terms for distributions of T with successively increasing skewness (β ranging from 5 to 10 in increments of 1, with the lowest skewness in the bottom line, increasingly for higher lines). As predicted, convergence is always faster when average health in the untreated sick state is higher. In Figure 5a (the ratio of 3rd to 2nd order terms) the ratio is about 0.32 for $\overline{H}_S = 0.1$, rapidly falling below 0.15 as $\overline{H}_S$ exceeds 0.2. Figure 5b similarly graphs the ratio of the 4th order (kurtosis-related) term to the 2nd order term. There, the lines reverse in sequence, with the smallest β parameters in the higher lines. The ratios are all below 0.75 and rapidly fall below 0.17 for values of $\overline{H}_S$ exceeding 0.2. These simulations demonstrate a reasonably rapid rate of convergence with these specific parameters, particularly when we limit the analysis to values of $\overline{H}_S \geq 0.2$.[12] Since QALY values below 0.2 correspond to near-death states, e.g., for a person with an untreatable and highly aggressive cancer, this is a helpful result.

Figure 5 also demonstrates that statistical moments produce uncertain effects on speed of convergence. For example, in the case of a four-parameter beta distribution, higher positive skew (larger values of β) results in slower convergence for a third-order Taylor expansion, but faster convergence for a fourth-order expansion. Nonetheless, our calculations illustrate how practitioners can readily calculate speed of convergence using their specific statistical moments and parameter values.

[12] The width of the support interval affects Figures 4a and 4b. As the support interval narrows for any 4-parameter Beta distribution, the variance falls with the square of that interval's width. Since Figure 4a contains σ^3 in the calculations of the 3rd moment (and similarly Figure 4b contains σ^4 to calculate the 4th moments), the ratios shown in these figures become smaller as the support interval narrows. The effect is linear in the support width in Figure 4a and quadratic in Figure 4b. The most pessimistic convergence would occur for support intervals near 1.0.

Sensitivity analyses (not shown) identify situations where the higher-order terms have ratios of the 3rd to 2nd order terms exceed 1.0, indicating slower convergence. The most important of these cases occurs when the two distributions have nearly-identical skewness, so that their difference is very small, and hence the ratio of the 2nd order to the 3rd order term relatively large. Even in these cases, it requires a relatively high degree of variance, which further increases the 3rd and 4th moment values. A similar phenomenon occurs when the two distributions have nearly-identical values of kurtosis.

Figure 5 here

VIII. CONCLUSIONS

Cost-effectiveness analysis has become one of the most successful economic methods in real-world applications to evaluate medical technologies, and its use continues to expand. However, standard cost-effectiveness frameworks fail to adequately account for the role of uncertainty in treatment effects. This can lead to misallocation of resources by health insurers and/or health care systems such as the British National Health Service. We develop a relatively straightforward and tractable way for analysts and real-world decision makers to account for treatment effect uncertainty as it affects risk-averse consumers.

Our analysis suggests the importance of measuring uncertainty in clinical trials. Variance is typically measured in clinical trials to establish the precision by which mean outcomes are known, but (to our knowledge) it never enters the actual calculation of the expected utility of health is affected by the treatment. We show that including estimates of variance and skewness (and potentially adding estimates of kurtosis) can improve accuracy of measures of value for medical technologies. Measuring skewness (and kurtosis) with meaningful accuracy may require increases in sample size in typical clinical trials used to establish treatment effects or meta-

analysis combinations of multiple trials on the same treatment. Assessing the extent of required sample size increases is best left to experts in that type of analysis, e.g., clinical trial biostatisticians.

Our analysis also illustrates the potential importance of accurate measures for relative risk-aversion, relative prudence (and perhaps relative temperance) over flows of health. For practical reasons, we employ the assumption that relative risk preferences are similar across health and non-health composite consumption, but this need not be so. Future research quantifying attitudes towards health risk can help sharpen our value assessments of medical technology. It seems plausible to assume that health risk attitudes will be constant across types of health interventions, suggesting that estimation need not be repeated for every individual health intervention.

Future cost-effectiveness analyses could implement our calculations using knowledge of the variance and skewness of treatment effects. Indeed, including measures of variance alone would improve the accuracy of value estimates, above and beyond the conventional approach of assuming risk-neutrality, but omitting the skewness term would penalize technologies with favorable changes in skewness in the distribution of health benefits.

Finally, we note that these results could help to guide design of new technologies in beneficial ways. Many diseases have ranges of consequences varying from relatively small to very large. Interventions that deal with the lower end of severity produce less total gain in expected utility than those treating patients with the most severe disease, even holding constant the magnitude of average improvement (μ_B).

Our analysis also directly answers a separate question raised by the ISPOR Task Force in 2018— whether there should be adjustments to value calculations based on the underlying

severity of disease. Our model shows that the proper answer is "yes," since introducing negative covariance between treatment outcomes and initial disease severity beneficially reduces the variance of disease outcomes. Put differently, this simply recognizes that humans experience diminishing marginal utility, so that utility gains arising from an improvement of some specific health gain are greater when they are bestowed on persons with lower initial health status.

Our model provides a systematic basis for incorporating measures of uncertainty in the outcomes of medical treatments that can differ importantly from traditional mean-based measures used in most cost-effectiveness analysis. These methods, grounded in economic theory and supported by literature-based estimates of relevant parameters, offer a tractable method for improving the accuracy of incremental cost-effectiveness analyses. Our simulations suggest that ignoring these stochastic components of treatment outcomes can seriously bias estimates of incremental cost-effectiveness, either in upward or downward direction, depending on how the new technology alters the risk profile of patient outcomes. Repairing this defect merely requires good estimates of skewness (and possibly kurtosis) measures of clinical outcomes in studies comparing medical interventions, and incorporation of those parameters into our new model.

BOX A

THREE WAYS TO INTERPRET PRUDENCE AND TEMPERANCE

Most economists are familiar with the standard measure of risk aversion (John W. Pratt, 1964, Kenneth J. Arrow 1963), but less so (perhaps wholly unfamiliar) with the higher-order terms of prudence and temperance [12, 13]. We offer several interpretations here.

1. Precautionary Savings.

In his pioneering work on the concept, Kimball defines prudence as "the sensitivity of a decision variable to risk" [14]. Define positive prudence as $\pi \equiv -\frac{u'''(Y_T)}{u''(Y_T)}$. An individual with positive prudence will respond to increases in the variance of future income by saving more today, known as "precautionary savings." Define the coefficient of absolute temperance, $\tau \equiv -\frac{u''''(Y_T)}{u'''(Y_T)}$ [14, 28]. Individuals with positive temperance seek to moderate their total exposure to risk [1, 28]. The prudence and temperance parameters—π and τ—are analogous to the more-familiar $r = -\frac{U''}{U'}$ except that they apply to the higher-order derivatives of the utility function. Positive prudence means that U''' > 0. Positive temperance means that U'''' < 0.

2. "Decreasing" behavior.

All utility functions with positive prudence ($\pi > 0$) have decreasing absolute risk aversion (DARA), a feature of utility functions widely accepted in the economics literature. Thus the following relationships generally hold: (1) positive risk aversion (r > 0) implies decreasing marginal utility of income; (2) positive prudence ($\pi > 0$) implies decreasing absolute risk aversion (DARA); (3) positive temperance ($\tau > 0$) implies decreasing DARA.

3. Mean-Preserving Spreads in Risk

Eeckhoudt et al (1995) explain the intuition behind the prudence and temperance terms [8]. It is well-understood that the degree of absolute risk-aversion measures a consumer's distaste for mean-preserving spreads in the distribution of consumption [41]. Risk-averse but prudent people dislike mean-preserving spreads. However, if they must accept one, they prefer that such spreads be applied to positive outcomes rather than negative ones. The degree of prudence represents the strength of their preferences in this respect [8].

Finally, temperate, prudent, and risk-averse people continue to dislike mean-preserving spreads, but prefer that spreads be applied to more positive outcomes. However, temperance means that such consumers derive diminishing marginal utility from successive rightward shifts of mean-preserving spreads of positive outcomes [8]. This is linked to kurtosis, which measures the degree to which positive skewness is produced by a few very extreme outliers or a larger number of more moderate outliers.

BOX B

FOUR-PARAMETER BETA DISTRIBUTIONS

First consider the (standardized) Beta distribution, with support over the [0,1] interval with parameters α, β, It describes proportions or percentages, with the following attributes:

$$\text{Mean} = \frac{\alpha}{\alpha+\beta}$$

$$Variance = \frac{\alpha\beta}{[(\alpha+\beta)^2(\alpha+\beta+1)]}$$

$$Pearson's\ Skewness = \frac{[2(\beta-\alpha)\sqrt{\alpha+\beta+1}]}{[(\alpha+\beta+2)\sqrt{\alpha\beta}]}$$

$$Pearson's\ Kurtosis = 3[\frac{(\alpha+\beta+1)(2(\alpha+\beta)^2+\alpha\beta(\alpha+\beta-6)]}{[\alpha\beta(\alpha+\beta+2)(\alpha+\beta+3)]}]$$

For any fixed ratio α/β, means remain unchanged, but variances shrink as the absolute parameter magnitudes increase. Thus, one can independently alter means and variances of Beta distributions. Beta distributions are positively skewed when $\beta > \alpha$, with mean < 0.5 and negatively skewed with mean > 0.5 when $\alpha > \beta$.

For our simulations, this unfortunately leads to adverse changes in skewness as mean treatment benefits increase, an undesirable feature since positive skewness is desirable in our model. Hence this distribution cannot serve our modeling goals. To resolve this, we use four-parameter Beta distributions, with a support over [c,d], a subset of [0,1]. Then we have:

$$Mean = c + (d-c)[\frac{\alpha}{\alpha+\beta}]$$

$$Variance = \frac{(d-c)^2\alpha\beta}{[(\alpha+\beta)^2(\alpha+\beta+1)]}$$

The distribution shifts and is "squeezed" into the [c,d] interval. The variance also changes by the factor $(d-c)^2$. Pearson's skewness and kurtosis remain unaffected. Thus, we can independently alter means without affecting skewness, and we can independently alter variances as previously noted. Thus, four-parameter Beta distributions are well-suited for our simulations. The only remaining defect is that we cannot independently change variance and skewness.

MATHEMATICAL APPENDIX

Here we prove Proposition 1 from the text.

Proposition 1: Suppose the health index has support on the interval $[0,1]$, and the utility function takes the form $U(c,H) = u(c)v(H)$, where v belongs to the class of HARA (hyperbolic absolute risk-aversion) utility functions. If H is a random variable with support $[0,1]$ and mean μ_0, then the Taylor expansion of $v(H)$ around μ_0 converges.

Proof: Any utility function in the HARA family can be written as $U(H) = \frac{1-\gamma}{\gamma}\left(\frac{\beta H}{1-\gamma}+\eta\right)^{\gamma}$ [27]. The Taylor expansion around the mean of H takes the form:

$$EU(H) \approx U(\mu_0) + U'(\mu_0)E(H-\mu_0)) + \frac{1}{2!}U''(\mu_0)E(H-\mu_0)^2 + \frac{1}{3!}U'''(\mu_0)E(H-\mu_0)^3 + \cdots$$

Since H has support on the unit interval, it is evident that $\lim_{n\to\infty}\frac{E(H-\mu_0)^{n+1}}{E(H-\mu_0)^n} = 0$. Moreover, defining the nth derivative of U as $U^{(n)}$, we can write:

$$\frac{U^{(n+1)}}{U^{(n)}} = \frac{\beta^{n+1}(\gamma-2)\ldots\frac{\gamma-n}{(1-\gamma)^{n-1}}\left(\frac{\beta\mu_0}{1-\gamma}+\eta\right)^{\gamma-(n+1)}}{\beta^{n}(\gamma-2)\ldots\frac{(\gamma-(n-1))}{(1-\gamma)^{n-2}}\left(\frac{\beta\mu_0}{1-\gamma}+\eta\right)^{\gamma-n}} = \frac{\beta(\gamma-n)}{\left(\frac{\beta\mu_0}{1-\gamma}+\eta\right)(1-\gamma)}$$

Thus, L'Hôpital's rule implies that $\lim_{n\to\infty}\frac{U^{(n+1)}}{(n+1)U^{(n)}} = \lim_{n\to\infty}\frac{\beta(\gamma-n)}{(n+1)\left(\frac{\beta\mu_0}{1-\gamma}+\eta\right)(1-\gamma)} = \frac{-\beta}{\left(\frac{\beta\mu_0}{1-\gamma}+\eta\right)(1-\gamma)}$.

Since β and γ are both finite scalars, since $0 \le \mu_0 \le 1$, and since $\lim_{n\to\infty}\frac{E(H-\mu_0)^{n+1}}{E(H-\mu_0)^n} = 0$, it follows that:

$$\lim_{n\to\infty}\frac{U^{(n)}E(H-\mu_0)^{n+1}}{(n+1)U^{(n+1)}E(H-\mu_0)^n} = 0$$

REFERENCES

1. Murphy, K.M. and R.H. Topel, "The Value of Health and Longevity." *Journal of Political Economy*, 2006. **114**(5): p. 871-904.
2. Rosen, S., "The Value of Changes in Life Expectancy." *Journal of Risk and Uncertainty*, 1988. **1**(3): p. 285-304.
3. Weinstein, M.C. and W.B. Stason, "Foundations of cost-effectiveness analysis for health and medical practices." *N Engl J Med*, 1977. **296**(13): p. 716-21.
4. Lising, A., A. Rosner, J. Gladman, and M. Drummond, "Payers' Use of ICER Reports in Decision Making," in *Association of Managed Care Pharmacy Nexus*. 2016, Journal of Managed Care and Specialty Pharmacy: National Harbor, Maryland.
5. Drummond, M.F., M.J. Sculpher, G.W. Torrance, B.J. O'Brien, and G.L. Stoddart, *Methods for the Economic Evaluation of Health Care Programmes*. Third Edition ed. 2005.
6. Lakdawalla, D.N., A. Malani, and J. Reif, "The insurance value of medical innovation." *Journal of Public Economics*, 2017. **145**: p. 94-102.
7. Lakdawalla, D.N., J.A. Romley, Y. Sanchez, J.R. Maclean, J.R. Penrod, and T. Philipson, "How cancer patients value hope and the implications for cost-effectiveness assessments of high-cost cancer therapies." *Health Aff (Millwood)*, 2012. **31**(4): p. 676-82.
8. Eeckhoudt, L., C. Gollier, and T. Schneider, "Risk-aversion, prudence and temperance: A unified approach." *Economics letters*, 1995. **48**(3-4): p. 331-336.
9. Lakdawalla, D.N., J.A. Doshi, L.P. Garrison, Jr., C.E. Phelps, A. Basu, and P.M. Danzon, "Defining Elements of Value in Health Care-A Health Economics Approach: An ISPOR Special Task Force Report [3]." *Value Health*, 2018. **21**(2): p. 131-139.
10. Garber, A.M. and C.E. Phelps, "Economic foundations of cost-effectiveness analysis." *J Health Econ*, 1997. **16**(1): p. 1-31.
11. Markowitz, H., "Portfolio selection." *The journal of finance*, 1952. **7**(1): p. 77-91.
12. Pratt, J.W., "Risk Aversion in the Small and in the Large." *Econometrica*, 1964. **32**(1): p. 122-36.
13. Arrow, K.J., "Liquidity preference." *Lecture VI in lecture notes for Economics*, 1963. **285**: p. 33-53.
14. Kimball, M.S., "Precautionary Saving in the Small and in the Large." *Econometrica*, 1990. **58**(1): p. 53-73.
15. Phelps, C.E. and G. Madhavan, "Using multicriteria approaches to assess the value of health care." *Value in Health*, 2017. **20**(2): p. 251-255.
16. Alatorre, C., G. Carter, C. Chen, C. Villarivera, V. Zarotsky, R. Cantrell, I. Goetz, R. Paczkowski, and D. Buesching, "A comprehensive review of predictive and prognostic composite factors implicated in the heterogeneity of treatment response and outcome across disease areas." *International journal of clinical practice*, 2011. **65**(8): p. 831-847.
17. Carter, G.C., R.A. Cantrell, V. Zarotsky, V.S. Haynes, G. Phillips, C.I. Alatorre, I. Goetz, R. Paczkowski, and L.B. Marangell, "Comprehensive review of factors implicated in the heterogeneity of response in depression." *Depression and anxiety*, 2012. **29**(4): p. 340-354.
18. Goetz, I., G.C. Carter, M. Lucero, V. Zarotsky, C.I. Alatorre, R.A. Cantrell, R. Paczkowski, and K.L. Sterling, "Review of treatment response in rheumatoid arthritis:

assessment of heterogeneity." *Current medical research and opinion*, 2011. **27**(4): p. 697-711.
19. Cantrell, R.A., C.I. Alatorre, E.J. Davis, V. Zarotsky, E. Le Nestour, G.C. Carter, I. Goetz, R. Paczkowski, and J. Sierra-Johnson, "A review of treatment response in type 2 diabetes: assessing the role of patient heterogeneity." *Diabetes, Obesity and Metabolism*, 2010. **12**(10): p. 845-857.
20. Yu, Y. and J. Cui, "Present and future of cancer immunotherapy: A tumor microenvironmental perspective." *Oncology letters*, 2018. **16**(4): p. 4105-4113.
21. Phelps, C., "A New Method to Determine the Optimal Willingness to Pay in Cost-Effectiveness Analysis." *Value in Health*, July, 2019. 22(7):785-791.
22. Markowitz, H., "Mean–variance approximations to expected utility." *European Journal of Operational Research*, 2014. **234**(2): p. 346-355.
23. Loistl, O., "The erroneous approximation of expected utility by means of a Taylor's series expansion: analytic and computational results." *The American Economic Review*, 1976. **66**(5): p. 904-910.
24. Hlawitschka, W., "The empirical nature of Taylor-series approximations to expected utility." *The American Economic Review*, 1994. **84**(3): p. 713-719.
25. Phelps, C.E., D.N. Lakdawalla, A. Basu, M.F. Drummond, A. Towse, and P.M. Danzon, "Approaches to Aggregation and Decision Making-A Health Economics Approach: An ISPOR Special Task Force Report [5]." *Value Health*, 2018. **21**(2): p. 146-154.
26. Basu, A. and S.D. Sullivan, "Toward a hedonic value framework in health care." *Value in Health*, 2017. **20**(2): p. 261-265.
27. Merton, R.C., "Optimum Consumption and Portfolio Rules in a Continuous-Time Model." *Journal of Economic Theory*, 1971. **3**: p. 373-413.
28. Kimball, M.S., "Precautionary Motives for Holding Assets," in *The New Palgrave Dictionary of Money and Finance*, P. Newman, M. Milgate, and J. Eatwell, Editors. 1992, Macmillan: London.
29. Ehrlich, I. and G.S. Becker, "Market Insurance, Self-Insurance, and Self-Protection." *The Journal of Political Economy*, 1972. **80**(4): p. 623-648.
30. Shafrin, J., T.T. Schwartz, T. Okoro, and J.A. Romley, "Patient Versus Physician Valuation of Durable Survival Gains: Implications for Value Framework Assessments." *Value Health*, 2017. **20**(2): p. 217-223.
31. Lakdawalla, D.N., T.J. Philipson, Y. Sanchez, R. Maclean, J.R. Penrod, and J. Romley, "The Value of Hope in Cancer Therapy." 2011, University of Southern California.
32. Noussair, C.N., S.T. Trautmann, and G. Van de Kuilen, "Higher order risk attitudes, demographics, and financial decisions." *Review of Economic Studies*, 2013. **81**(1): p. 325-355.
33. Malani, A. and T.J. Philipson, "Can medical progress be sustained? Implications of the link between development and output markets." 2011, Bureau of Economic Research Inc NBER Working Papers: 17011 2011.
34. Phelps, C.E. and G. Madhavan, "Adverse Events from Medical Treatments (AEMTs) in Healthcare Technology Assessment (HTA)." 2019, University of Rochester.
35. Yaari, M.E., "Uncertain lifetime, life insurance, and the theory of the consumer." *The Review of Economic Studies*, 1965. **32**(2): p. 137-150.
36. Bommier, A. and B. Villeneuve, "Risk aversion and the value of risk to life." *Journal of Risk and Insurance*, 2012. **79**(1): p. 77-104.

37. Bode, B.W., M.A. Testa, M. Magwire, P.M. Hale, M. Hammer, L. Blonde, and A. Garber, “Patient-reported outcomes following treatment with the human GLP-1 analogue liraglutide or glimepiride in monotherapy: results from a randomized controlled trial in patients with type 2 diabetes.” *Diabetes, Obesity and Metabolism*, 2010. **9999**(999A).
38. Christelis, D., D. Georgarakos, T. Jappelli, and M.v. Rooij, “Consumption Uncertainty and Precautionary Saving.” *The Review of Economics and Statistics*, 2019. **Forthcoming**: p. 1-45.
39. Arrow, K.J., *Essays in the theory of risk-bearing*. Markham Economics Series. 1971, Chicago: Markham Publishing Company.
40. Romley, J.A., Y. Sanchez, J.R. Penrod, and D.P. Goldman, “Survey results show that adults are willing to pay higher insurance premiums for generous coverage of specialty drugs.” *Health Affairs*, 2012. **31**(4): p. 683-90.
41. Rothschild, M. and J.E. Stiglitz, “Increasing risk: I. A definition.” *Journal of Economic theory*, 1970. **2**(3): p. 225-243.

TABLES AND FIGURES

Table 1. Summary of variance and skewness effects on certainty-equivalence ratio.

	$\Delta\sigma^2 > 0$	$\Delta\sigma^2 < 0$
$\Delta\gamma_1 > 0$	Bad, but muted, possible reversal	Good, amplified. Both terms increase ϵ_Δ
$\Delta\gamma_1 < 0$	Bad; both terms diminish ϵ_Δ	Good but muted, possible reversal

Figure 1. Contribution of variance, skewness, and treatment effects to the certainty-equivalence ratio.

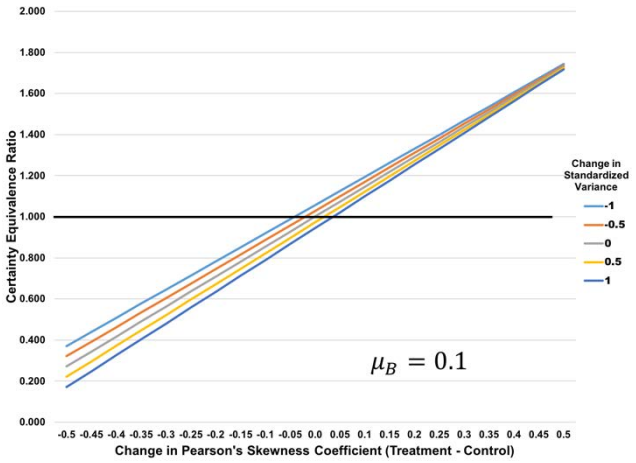

Figure 1(a)

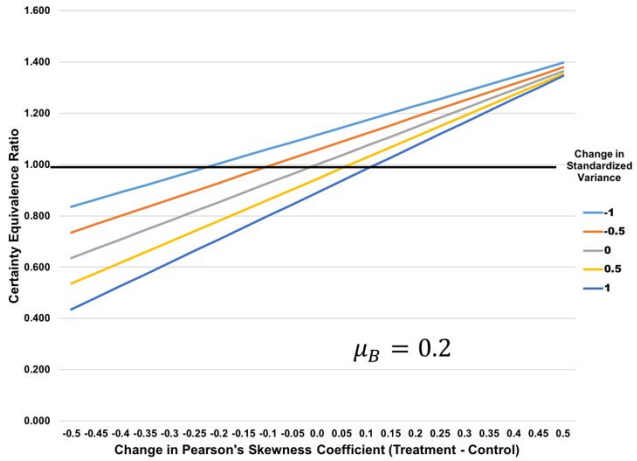

Figure 1(b)

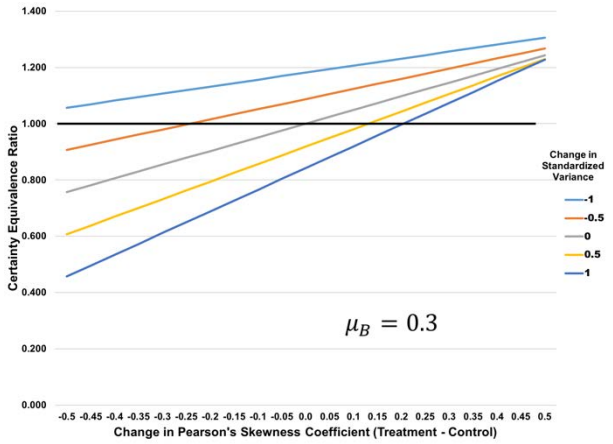

Figure 1(c)

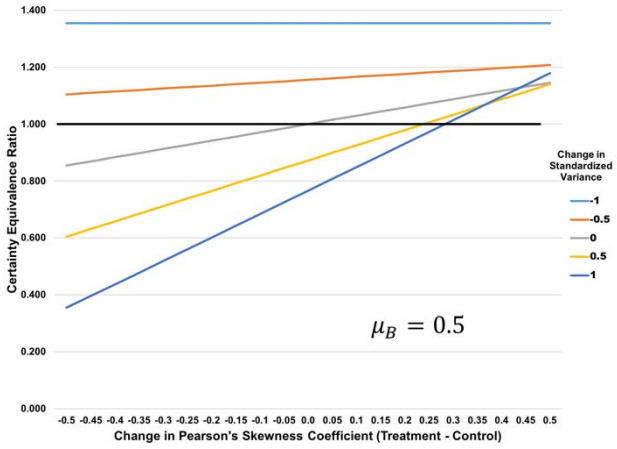

Figure 1(d)

Notes: The vertical axis shows the value of the certainty-equivalence ratio, ϵ_Δ. The five lines represent different values for the change in standardized variance between the treatment and control, defined as $\left(\frac{\Delta\sigma^2}{\mu_B^2}\right)$. The upper line represents a decline of 1 (-1 value), the lowest line an increase of 1 (+1 value), the middle line no change in variance (0 value), and the other two halfway between the middle and the extremes (-0.5 or + 0.5). The horizontal axis shows the difference in Pearson's skewness coefficient, γ_1, between the treatment and control. The figures show increasing degrees of health gain from the treatment (T) compared with the untreated (or comparison treatment) state, with values of μ_B ranging from 0.1 to 0.5. In all cases, average health in the untreated state is $\overline{H_S} = 0.5$. The heavy line at $\epsilon_\Delta = 1$ represents the breakeven point where the traditional mean-based estimate of value equals the stochastic-based estimate. These demonstrate the effects of μ_B on the estimated values of ϵ_Δ for given values of $\overline{H_S}$.

Figure 2. Baseline disease severity and the certainty-equivalence ratio.

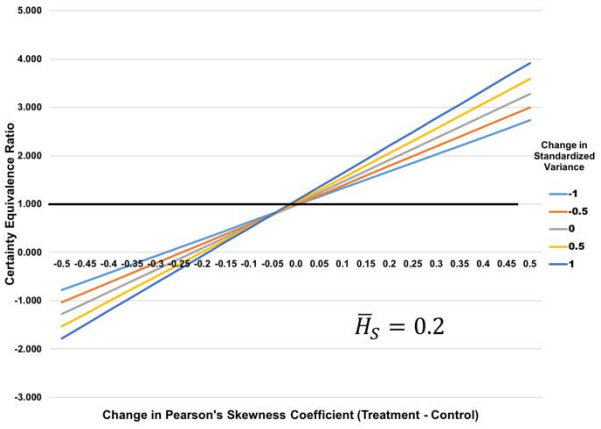

Figure 2(a)

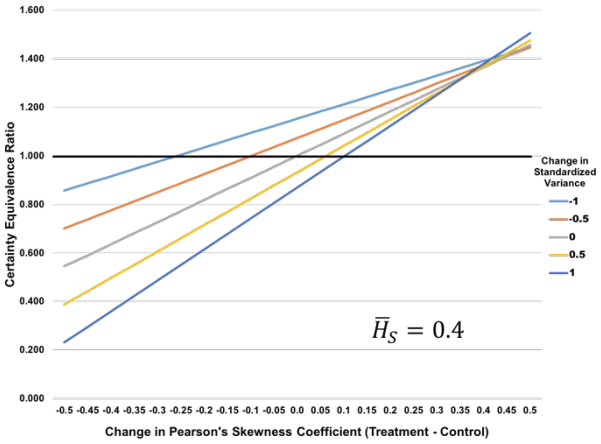

Figure 2(b)

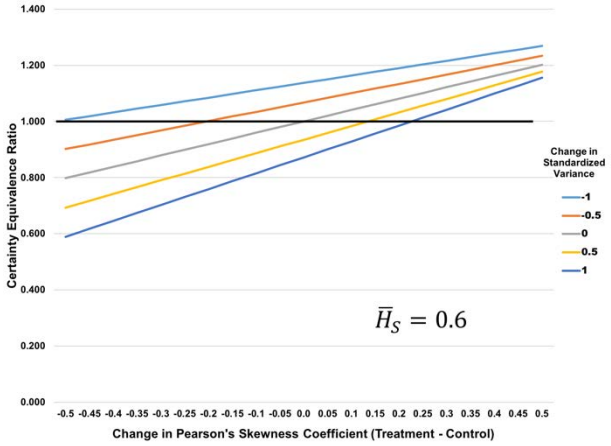

Figure 2(c)

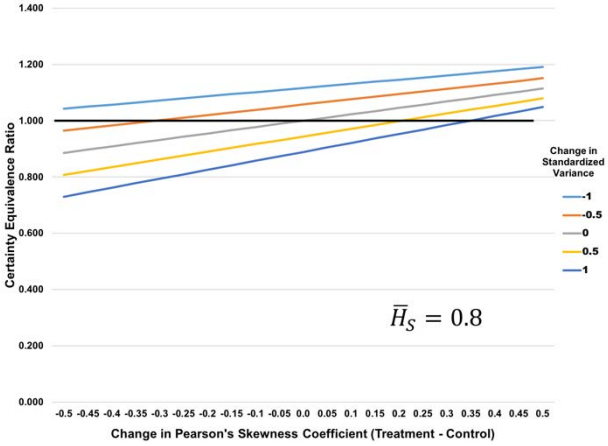

Figure 2(d)

Notes: The structure of these figures is identical to that in Figure 1 except that they hold $\mu_B = 0.25$ and vary $\overline{H}_S$ from 0.2 to 0.8. These demonstrate the effects of $\overline{H}_S$ on the estimated values of the certainty-equivalence ratio, ϵ_Δ, for given values of μ_B.

Figure 3. Probability densities of treatment effects and implications for certainty-equivalence ratio.

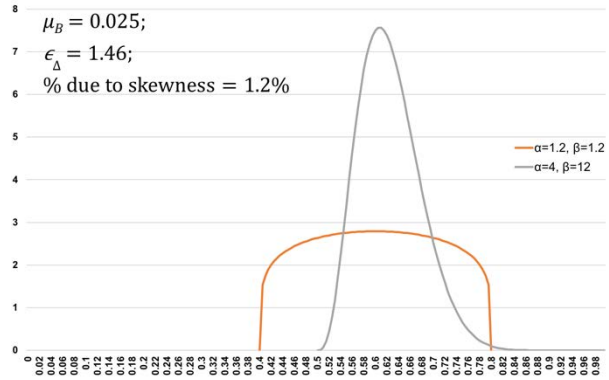

Figure 3(a)

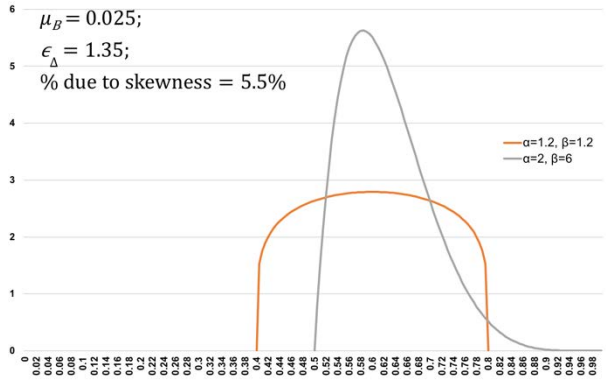

Figure 3(b)

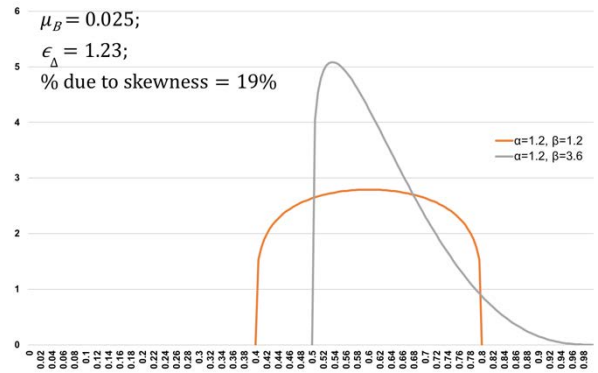

Figure 3(c)

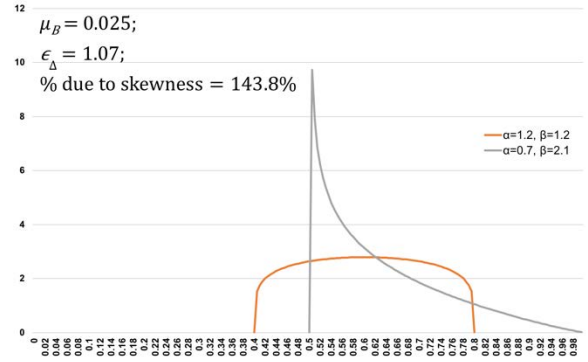

Figure 3(d)

Notes: These Figures employ four-parameter Beta distributions (see "BOX B" in main text for details). In all panels, the Sick distribution has support on the [0.4, 0.8] interval with parameters $\alpha = 1.2,\ \beta = 1.2$, so it has no skewness. The distribution of outcomes for T all have a ratio of $\alpha{:}\beta$ of 1/3, with increasingly smaller values, so that the variance and skewness both increase as one moves from panels (a) to (d). The increasing variance of T causes the certainty-equivalence ratio, ϵ_Δ, to decline from 1.46 to 1.07. In panel (d), the variance of the T distribution exceeds that of the S distribution, but the positive skewness still leaves ϵ_Δ = 1.07, demonstrating the value of skewness even when the T distribution has worse variance than the S distribution.

Figure 4. Treatment effect sizes and implications for the certainty-equivalence ratio.

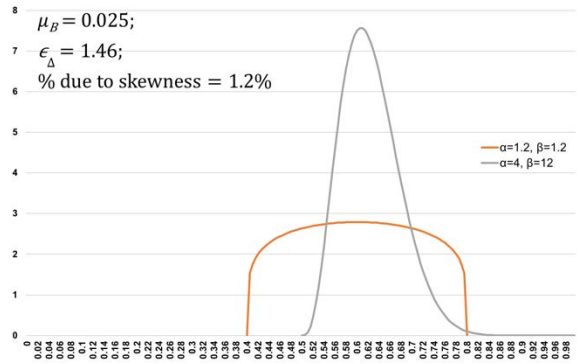

Figure 4(a)

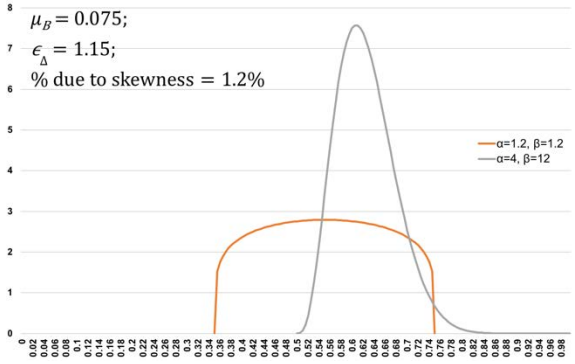

Figure 4(b)

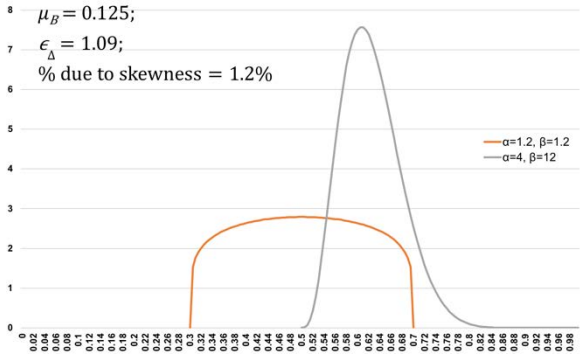

Figure 4(c)

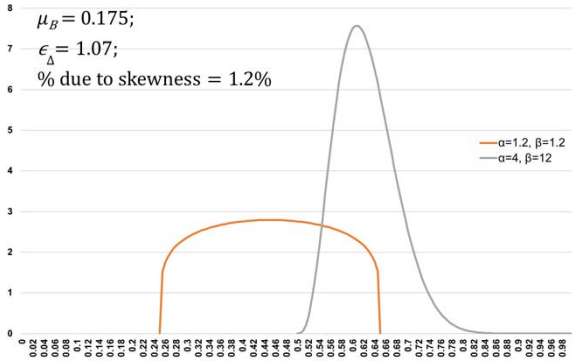

Figure 4(d)

Notes: Figure 4(a) is identical to Figure 3(a), with a "moderate" degree of positive skewness in T. Panels 4(b) to 4(d) successively shift the distribution of S leftward, thus increasing μ_B from 0.025 to 0.175. As μ_B increases, the importance of the stochastic components falls, so the certainty-equivalence ratio, ϵ_Δ, declines from 1.46 to 1.07 when moving from Panels (a) to (d).

Figure 5. Speed of convergence in Taylor Series approximation to value of medical technology.

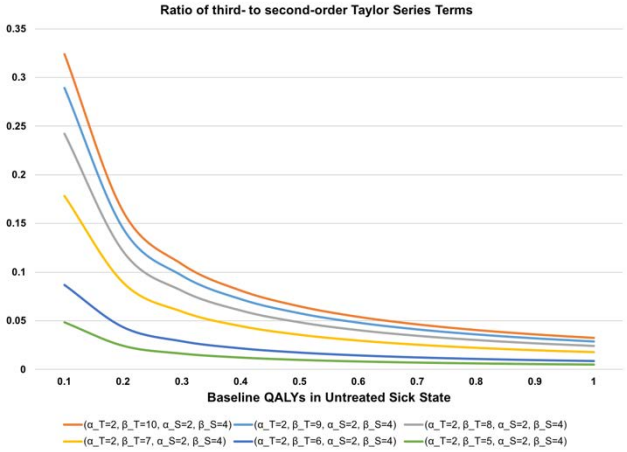

Figure 5(a)

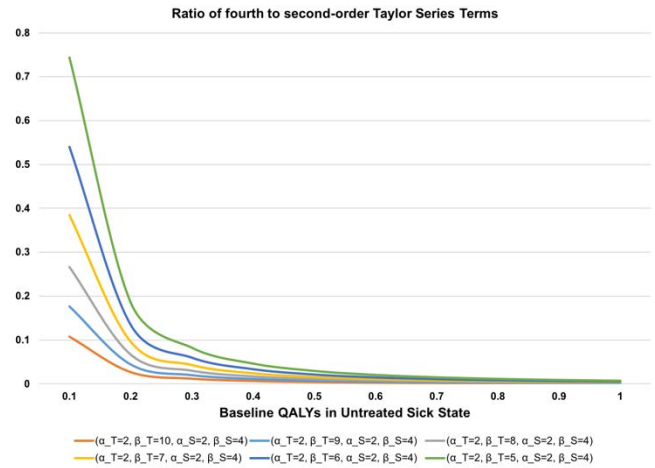

Figure 5(b)

Notes: These figures demonstrate the rate of convergence of the Taylor series by taking the ratio of the 3rd to the 2nd order terms (Figure 5(a)) and the 4th to 2nd order terms (Figure 5(b)). Smaller ratios demonstrate faster convergence. In both panels, the S outcomes have Beta distribution parameters of $\alpha = 2; \beta = 4$. In Panel 5(a), the uppermost line has the greatest skewness ($\alpha = 2; \beta = 10$), with the value of β declining (serially) in the lower lines by unit values to 9, … 5, holding α constant at 2. Therefore, the skewness declines as one moves from the uppermost to the lowermost lines. In Figure 5(b), the sequence is reversed (following the formula for kurtosis in the 4-parameter Beta distribution). Otherwise Figures 5(a) and 5(b) have similar structure.