DISCUSSION PAPER NO. 43

Point Estimation and Risk Preferences

by

David P. Baron

March 1973

Forthcoming in the *Journal of the American Statistical Association*

Point Estimation and Risk Preferences

David P. Baron*
Northwestern University

1. Introduction

The decision-theoretic approach to point estimation involves the choice of an estimate or decision to maximize the expected utility, or equivalently to minimize the expected loss, associated with the estimate. The expected utility or loss depends on both the payoff function and the risk preferences of the decision maker and even though the payoff function associated with an estimate or a decision may be invariant among decision makers, preferences for alternative payoffs are likely to differ. LaValle [4] has argued that it is important to individually consider both the reward function and the risk preference or utility function rather than to treat both as a single composite function. The purpose of this paper is to indicate the influence of risk aversion on point estimates for classes of payoff functions including the piecewise linear and quadratic payoff functions. For example, with a quadratic payoff function risk aversion results in optimal point estimates that are between zero and the mean of the population being estimated, while with a piecewise linear utility function the optimal point estimate is less if the utility function is concave than if the utility function is linear.

*This research was supported in part by the National Science Foundation. The author wishes to thank Wei Chong for performing the computations and Robert L. Winkler and the referees for their helpful comments.

The point of view adopted in this paper is that if a decision is associated with a point estimate the proper approach is to assess the consequences or payoffs associated with the estimate and to choose the estimate so as to maximize the expected utility of the payoffs or equivalently to minimize the expected loss. If the point estimate is to be used solely for inferential purposes, such as judging the scientific truth or falsity of a proposition, or is to be reported for whatever purposes others may wish to use it, the statistician should report the likelihood function associated with the experiment. As suggested by Hildreth [3] the statistician may also wish to report "solutions to representative decision problems" as well as the estimates based on classical properties such as unbiasedness, consistency, efficiency, etc. The results of this paper are relevant in the former context in which a known decision problem is at hand and in the latter context in which the statistician wishes to report solutions to decision problems for the benefit of unknown or remote clients.

A point estimation problem involves the choice of an estimate (a) from a set A of possible estimates, where the consequences or payoffs associated with the estimate depend on an element θ of a space Θ. The element θ may be a parameter of the distribution of a random variable or may be the random variable itself. The distribution function of θ is denoted by $P(\theta)$ and may be thought of as either 1) a subjective prior distribution based upon no sample data, 2) a distribution based solely on sample data, or 3) a posterior distribution incorporating both prior and sample information. The payoff is denoted by $g(\theta,a)$ where g is defined on $\Theta \times A$ and takes on values on the real line. The risk preferences of the decision maker are assumed to be represented by a strictly increasing, concave, twice differentiable, utility function U defined over the set of possible values for $g(\theta,a)$. The statistician chooses the estimate (a) so as to attain the

$$\max_{a \in A} \int_{\Theta} U(g(\theta,a))dP(\theta). \tag{1.1}$$

This is equivalent to defining a nonnegative loss function[1]

$$L(\theta,a) = -U(g(\theta,a)) + \max_{a \in A} U(g(\theta,a)) \tag{1.2}$$

and choosing the estimate (a) so as to attain the[2]

$$\min_{a \in A} \int_{\Theta} L(\theta,a)dP(\theta). \tag{1.3}$$

In a more general setting LaValle [4] has considered the effects of risk preferences on the certainty equivalent for decision problems such as (1.1) (and hence (1.3)). The emphasis in this paper is on the impact of risk preferences on the optimal point estimate.

Risk preferences will be measured by the Arrow-Pratt [1,6] index of absolute risk aversion $r_U(y)$ defined by $r_U(y) = - U''(y)/U'(y)$, where ($'$) and ($''$) denote first and second derivatives, respectively. Pratt has demonstrated that an increase in $r_U(y)$ for all y results in an increase in the risk premium, where the risk premium $\pi(a,P)$ for a given estimate (a) is defined as

$$\int_{\Theta} U(g(\theta,a))dP(\theta) \equiv U\left(\int_{\Theta} g(\theta,a)dP(\theta) - \pi(a,P)\right).$$

For U strictly concave $\pi(a,P) > 0$ indicating that the certainty equivalent $(\int_{\Theta} g(\theta,a)dP(\theta) - \pi(a,P))$ is less than the expected value of the payoff. For a fixed (a) Pratt's result implies that the certainty equivalent is decreasing in absolute risk aversion, and LaValle [4] has shown that this result also holds when the optimal $a \in A$ is chosen.

If $r_U(y)$ is decreasing (increasing)(constant) in y, U is said to exhibit decreasing (increasing)(constant) absolute risk aversion. Thus, a positive translation in $g(\theta,a)$ results in a decrease (increase) (no change) in the risk premium of $g(\theta,a)$ if the utility function exhibits decreasing (increasing)(constant) absolute risk aversion.[3] A utility function U is said to exhibit increasing (decreasing)(constant) proportional or relative risk aversion if

$$r_U^*(y) \equiv y r_U(y) = -\,(U''(y)\cdot y/U'(y))$$

is increasing (decreasing)(constant) in y. The interpretation of relative risk aversion is that the decision maker is indifferent between a risk $yg(\theta,a)$ and $E(yg(\theta,a)) - \pi(a,P) = E(yg(\theta,a)) - y\pi^*(a,P)$ for certain where E denotes expectation and $\pi^*(a,P) = \frac{1}{y}\pi(a,P)$ is the relative risk premium.[4]

To investigate the implications of risk aversion for optimal point estimates, it is sufficient to work within the context of expected utility maximization, but the loss function formulation will be used in order to provide an interpretation of the results in the context of classical estimation procedures. To illustrate the effects of risk aversion, two common estimation problems developed from the newsboy or piecewise-linear payoff function and the quadratic payoff function will be considered. Theoretical results are presented in the next two sections, and numerical results are given in Section 4.

2. Quadratic Payoff Function

The most commonly used loss function in point estimation is the squared-error or quadratic function $k(\theta-a)^2$ defined on $\Theta \subset A = R = (-\infty,+\infty)$. With a linear utility function, a quadratic loss function may be obtained from the payoff function

$$g(\theta,a) = 2k(\theta a - \tfrac{1}{2} a^2), \; k > 0. \tag{2.1}$$

Such quadratic functions are frequently used as approximations of other more complex functions and arise in economics when linear demand and marginal cost functions are used. For example, consider a price-taking firm in a competitive market in which the price p is uncertain, and let the total cost function C(a) for a quantity (a) be $C(a) = ba + ca^2$, $c > 0$, $b > 0$. The profit of the firm is $(p-b)a - ca^2$, which with a transformation of variables is of the form of (2.1).[5]

The general loss function for the quadratic payoff function in (2.1) is

$$L(\theta,a) = U(k\theta^2) - U(2k(\theta a - \tfrac{1}{2} a^2)). \tag{2.2}$$

If the utility function is linear, the loss function is

$$L(\theta,a) = k(\theta-a)^2 , \tag{2.3}$$

and the optimal point estimate is the mean of the distribution of θ. The effect of risk aversion on the optimal point estimate is to move it towards zero if the optimal estimator is nonzero. A more general theorem first will be proved for a class of payoff functions that includes (2.1), and then the result will be interpreted in terms of the classical properties used to evaluate point estimators.

Theorem 1: Let $U_1(y)$ and $U_2(y)$ be two utility functions such that $r_{U_1}(y) \geq r_{U_2}(y) \geq 0$ for all y and $r_{U_1}(y) > r_{U_2}(y)$ for some $y = g(\theta,\hat{a}_2)$ (or on some subinterval containing y) with positive probability, and let $\hat{a}_1$ be optimal for U_1 and $\hat{a}_2$ (finite) be optimal for U_2. Let $g(\theta,a)$ be concave and continuously differentiable in a. A) If $g(\theta,a)$ is monotone increasing (decreasing) in θ and $\frac{\partial g(\theta,a)}{\partial a} = g'(\theta,a)$ is monotone increasing (decreasing) in θ for the optimal $\hat{a}_2$, then $\hat{a}_1 < \hat{a}_2$. B) If $g(\theta,a)$ is monotone increasing (decreasing) in θ and $g'(\theta,a)$ is monotone decreasing (increasing) in θ for the optimal $\hat{a}_2$, then $\hat{a}_1 > \hat{a}_2$. C) If $g(\theta,\hat{a}_2)$ is constant for all $\theta \in \Theta$, then $\hat{a}_1 = \hat{a}_2$. D) If $g(\theta,a=0) = 0$ for all $\theta \in \Theta$, and $\max_{a \in A} \int_\Theta g(\theta,a)dP(\theta) = 0$, then for all U concave $\hat{a} = 0$ is optimal.

Proof: A) The necessary optimality condition for the expected loss in (2.2) with U_2 is

$$- \int_{-\infty}^{\infty} U_2'(g(\theta,\hat{a}_2))\, g'(\theta,\hat{a}_2)dP(\theta) = 0. \tag{2.4}$$

For U_2 concave the estimator satisfying (2.4) is a unique global optimum. For $\hat{a}_2$ finite, there exists a θ^* satisfying $g'(\theta^*,\hat{a}_2) = 0$ or else (2.4) would not hold and $\hat{a}_2$ would not be finite. Dividing (2.4) by $U_2'(g(\theta^*,\hat{a}_2))$ and rewriting yields

$$- \int_{-\infty}^{\theta^*} \frac{U_2'(g(\theta,\hat{a}_2))}{U_2'(g(\theta^*,\hat{a}_2))} g'(\theta,\hat{a}_2)dP(\theta) - \int_{\theta^*}^{\infty} \frac{U_2'(g(\theta,\hat{a}_2))}{U_2'(g(\theta^*,\hat{a}_2))} g'(\theta,\hat{a}_2)dP(\theta) = 0. \tag{2.5}$$

Doing the same for U_1 evaluated at $\hat{a}_2$ and subtracting from (2.5) yields

$$-\int_{-\infty}^{\theta^*}\left[\frac{U_2'(g(\theta,\hat{a}_2))}{U_2'(g(\theta^*,\hat{a}_2))} - \frac{U_1'(g(\theta,\hat{a}_2))}{U_1'(g(\theta^*,\hat{a}_2))}\right] g'(\theta,\hat{a}_2)dP(\theta)$$

$$-\int_{\theta^*}^{\infty}\left[\frac{U_2'(g(\theta,\hat{a}_2))}{U_2'(g(\theta^*,\hat{a}_2))} - \frac{U_1'(g(\theta,\hat{a}_2))}{U_1'(g(\theta^*,\hat{a}_2))}\right] g'(\theta,\hat{a}_2)dP(\theta) \qquad (2.6)$$

Pratt [6,Eq.20] indicates that $r_{U_1}(y) \geq r_{U_2}(y)$ for all y and $r_{U_1}(y^*) > r_{U_2}(y^*)$ is equivalent to

$$(U_1'(y^*)/U_1'(w)) > (U_2'(y^*)/U_2'(w)),\ y^* < w. \qquad (2.7)$$

Under the assumptions of A) for $\theta < \theta^*$, the term in brackets in the first integral in (2.6) is negative (positive) from (2.7), since $g(\theta,\hat{a}_2)$ is monotone increasing (decreasing) in θ. Since $g'(\theta,\hat{a}_2)$ is monotone increasing (decreasing), $g'(\theta,\hat{a}_2) < (>)\ 0$ for $\theta < \theta^*$, so the first integral is positive. The second integral is also positive using the inverse of (2.7) and noting that $g'(\theta,\hat{a}_2) > (<)\ 0$ for $\theta > \theta^*$, so the expression in (2.6) is negative. From (2.4) the expression in (2.6) is

$$-\int_{-\infty}^{\infty} U_1'(g(\theta,\hat{a}_2))\ g'(\theta,\hat{a}_2)dP(\theta) > 0. \qquad (2.8)$$

The strict convexity of expected loss implies that the optimal point estimate must be decreased from $\hat{a}_2$ for (2.8) to hold as an equality so $\hat{a}_1 < \hat{a}_2$.

B) Arguing as above under the assumptions of B) indicates that the term on the left-hand side of (2.8) is negative, so $\hat{a}_1 > \hat{a}_2$.

C) Under the assumptions of C) the expressions in brackets in (2.6) equal zero, and consequently, $\hat{a}_2 = \hat{a}_1$.

D) For U concave $\int_\Theta U(g(\theta,a))dP(\theta) = U(\int_\Theta g(\theta,a)dP(\theta)-\pi(a,P)) \leq U(\int_\Theta g(\theta,a)dP(\theta))$. Then under the assumptions of D), since minimizing the expected loss is equivalent to maximizing expected utility, the optimal estimate is such that $\max_{a\in A} U(\int_\Theta g(\theta,a)dP(\theta) - \pi(a,P)) = U(\max_{a\in A} (\int_\Theta g(\theta,a)dP(\theta)-\pi(a,P))) \leq U(\max_{a\in A} \int_\Theta g(\theta,a)dP(\theta)) = U(0)$. Since $U(0)$ may be attained for $a = 0$, $\hat{a} = 0$ is optimal for U concave. //

The quadratic payoff function in (2.1) satisfies assumption A) for $\hat{a}\in(0,+\infty)$, B) for $\hat{a}\in(-\infty,0)$, C) for $\hat{a} = 0$ and D) for $\int_\Theta \theta dP(\theta) = 0$. The following corollary specializes Theorem 1 for the quadratic payoff function in (2.1).

Corollary 1: Let $U_1(y)$ be at least as absolute risk averse as $U_2(y)$ for all y and more risk averse for some $y = g(\theta,\hat{a}_2)$ (or on some sub-interval) with positive probability, and let $\hat{a}_1$ be optimal for U_1 and $\hat{a}_2$ be optimal for U_2. Then, $\hat{a}_1 < \hat{a}_2$ if $\hat{a}_2 > 0$; $\hat{a}_1 > \hat{a}_2$ if $\hat{a}_2 < 0$; $\hat{a}_1 = \hat{a}_2 = 0$ if $\hat{a} = 0$ is optimal for some concave utility function, and $\int_\Theta \theta\, dP(\theta) = 0$ implies $\hat{a} = 0$ for all concave utility functions.

The following corollary specializes Theorem 1 for utility functions with decreasing (increasing)(constant) absolute risk aversion.

Corollary 2: The optimal estimate for the loss function in (2.2) is an increasing (decreasing)(constant) function of y for $\hat{a} > 0$ (and vice versa for $\hat{a} < 0$) if $U(y+g(\theta,a))$ exhibits decreasing (increasing)(constant) absolute risk aversion. If $\hat{a} = 0$, the optimal point estimate is independent of y.

Corollary 2 follows from Theorem 1 by letting $U_1(y) = U(y)$ and $U_2(y) = U(y+\Delta y)$, $\Delta y > 0$. If $g(\theta,a)$ is measured in monetary units, y may represent the wealth of the decision maker.

With the quadratic payoff function increased risk aversion moves the optimal estimate away from the mean and towards zero. The latter occurs because as the point estimate moves closer to zero the "risk" involved in the estimate decreases. To examine this "risk," consider the following example involving the payoff function in (2.1) and a constant absolutely risk averse utility function $U(y) = -\exp(-\alpha y)$, $\alpha > 0$. The parameter θ may be interpreted as, for example, a random variable such as an uncertain price or as another example, may be interpreted as the unknown mean of a population with a known variance σ^2. For the latter case θ is considered to be a random variable, and $P(\theta)$ is a posterior distribution developed from an informationless prior. If a random sample of size n (large) is taken from the population, $P(\theta)$ is approximately normal with mean $\bar{x}$ and variance $\frac{\sigma^2}{n}$. Then the expected utility, upon completing the square in the exponent and integrating, is

$$\int_{-\infty}^{+\infty} U(g(\theta,a))dP(\theta) = -\exp(-\alpha(2k(\bar{x}a-(1/2)a^2) - 2\alpha k^2 a^2 \sigma^2/n)).$$

The risk premium is $2\alpha k^2 a^2 \sigma^2/n$, which is increasing in the index of absolute risk aversion α and increases as the distance from (a) to the origin increases. The optimal estimator is $\hat{a} = \frac{\bar{x}}{1+2\alpha k\sigma^2/n}$ which for $\mu > (<)\ 0$ is decreasing (increasing) in α and is less (greater) than the mean $\bar{x}$. As the index of absolute risk aversion increases, the decision maker reduces (increases) the estimate from the mean for $\bar{x} > (<)\ 0$ in order to increase expected utility by increasing the certainty equivalent $(2k(\bar{x}\hat{a}-(1/2)\hat{a}^2)-2\alpha k^2\hat{a}^2\sigma^2/n)$. The "bias" (in the classical sense) of the optimal point estimator is $\left(\frac{\bar{x}}{1+2\alpha k\sigma^2/n} - \bar{x}\right) = -\frac{2\alpha\bar{x}k\sigma^2/n}{1+2\alpha k\sigma^2/n}$. This "bias" results from risk aversion and is not to be thought of as undesirable in a decision context, since optimality is defined in terms of the decision maker's utility function. The expected loss in payoff[6] in (2.3) resulting from the optimal decision $\hat{a}$ is

$$E_\theta L(\theta,\hat{a}) = E_\theta(k(\theta-\hat{a})^2) = k(\sigma^2+(\bar{x} - \frac{\bar{x}}{1+2\alpha k\sigma^2/n})^2),$$

where the term $(\bar{x} - \frac{\bar{x}}{1+2\alpha k\sigma^2/n})^2$ is the increase in the expected loss of payoff resulting from the "bias" in the estimation caused by risk aversion. If the estimate is to be used solely for inferential purposes, risk preferences should be suppressed eliminating this "bias." This corresponds to using a linear utility function ($\alpha=0$) yielding an expected loss in payoff of $k\sigma^2$.

For a strictly concave utility function in the limit as n increases the optimal point estimate approaches the sample mean, so the point estimate obtained with a strictly concave utility function is asymptotically unbiased. Both the "bias" and the variance (σ^2/n) go to zero as n increases, so the risk averse point estimator is consistent and is also asymptotically squared-error efficient. This same asymptotic behavior will result in general provided that in the limit the distribution $P(\theta)$ places mass one on the true value of the parameter.

The optimal point estimate for the above example is decreasing (increasing)(constant) in k for μ > (<)(=) 0 and $\alpha > 0$, and this result may be generalized using the notion of relative risk aversion.

Theorem 2: Let U be concave and $g(\theta,a)$ be as given in (2.1). A) If $\hat{a} = 0$, the optimal point estimate is constant in k. B) The optimal point estimate is increasing (decreasing)(constant) in k if 1) $\hat{a} < 0$ and U exhibits increasing (decreasing)(constant) relative risk aversion or if 2) $\hat{a} > 0$ and U exhibits decreasing (increasing)(constant) relative risk aversion.[7,8]

Proof: Implicit differentiation of the necessary optimality condition in (2.4) yields $\frac{d\hat{a}}{dk} = -\frac{1}{EL_{aa}}[-2\int_{-\infty}^{+\infty} U''(2k(\theta\hat{a}-(1/2)\hat{a}^2))(\theta-\hat{a})(\theta\hat{a} - (1/2)\hat{a}^2)dP(\theta)]$, where EL_{aa} is the (positive) second-derivative of the expected loss evaluated at $\hat{a}$.

A) If $\hat{a} = 0$, then $\theta\hat{a} - (1/2)\hat{a} = 0$ and $\frac{d\hat{a}}{dk} = 0$.

B) Let $\theta^* \equiv \hat{a}$ and consider the case in which $\hat{a} > 0$. For $\hat{a} > 0$, $r^*_U(y)$ increasing implies that

$$-\frac{U''(2k(\theta\hat{a}-(1/2)\hat{a}^2))}{U'(2k(\theta\hat{a}-(1/2)\hat{a}^2))}2k(\theta\hat{a}-(1/2)\hat{a}^2) \geq 2kr^*_U(2k(\theta^*\hat{a}-(1/2)\hat{a}^2)), \quad (2.9)$$

since for $\theta - \theta^* > 0$ and $\hat{a} > 0$, $2k(\theta\hat{a}-(1/2)\hat{a}^2) > 2k(\theta^*\hat{a}-(1/2)\hat{a})^2$. Multiply both sides of the inequality in (2.9) by $-U'(2k(\theta\hat{a}-(1/2)\hat{a}^2))(\theta-\hat{a})$ to obtain for $\theta - \hat{a} > 0$

$$U''(2k(\theta\hat{a}-(1/2)\hat{a}^2))\, 2k(\theta\hat{a}-(1/2)\hat{a}^2)(\theta-a)$$

$$\leq -2kr^*_U(2k(\theta^*\hat{a}-(1/2)\hat{a}^2))\, U'(2k(\theta\hat{a}-(1/2)\hat{a}^2))(\theta-\hat{a}). \quad (2.10)$$

For $\theta - \hat{a} < 0$ and $\hat{a} > 0$,

$$-\frac{U''(2k(\theta\hat{a}-(1/2)\hat{a}^2))}{U'(2k(\theta\hat{a}-(1/2)\hat{a}^2))}2k(\theta\hat{a}-(1/2)\hat{a}^2) \leq 2kr^*_U(2k(\theta^*\hat{a}-(1/2)\hat{a}^2)). \quad (2.11)$$

Multiplying both sides of the inequality in (2.11) by $[-U'(2k(\theta\hat{a}-(1/2)\hat{a}^2))(\theta-\hat{a})]$ indicates that (2.10) holds for all θ. Integrating and noting that the right-hand side of (2.10) is the necessary optimality condition multiplied by a constant implies that $\frac{d\hat{a}}{dk} < 0$. The opposite result obtains for $r^*_U(y)$ decreasing, and similar analysis indicates that for constant relative risk aversion $\frac{d\hat{a}}{dk} = 0$. Next consider $\hat{a} < 0$. Then $[\theta\hat{a}-(1/2)\hat{a}^2]$ is increasing in θ as $(\theta-\hat{a})$ decreases. The inequality in (2.10) thus is reversed for $\theta - \hat{a} > 0$ and the inequality in (2.11) is reversed for $\theta - \hat{a} < 0$, and hence (2.10) holds for all θ. This implies that $\frac{d\hat{a}}{dk} > (<)(=)\ 0$ for increasing (decreasing)(constant) relative risk aversion. //

The following table summarizes the results of Theorem 2.

	Sign of $\frac{d\hat{a}}{dk}$		
Relative Risk Aversion	$\hat{a} < 0$	$\hat{a} = 0$	$\hat{a} > 0$
increasing	+	0	-
constant	0	0	0
decreasing	-	0	+

The interpretation of the result of Theorem 2 is, for example, that with increasing relative risk aversion an increase in k results in a greater proportional risk (for $\hat{a} \neq 0$) and to reduce the risk the optimal estimate moves towards zero. The opposite effect occurs for decreasing relative risk aversion.

3. Piecewise Linear Payoff Function

The effect of risk aversion on the optimal point estimate depends importantly on the form of the payoff function. In this section the frequently used piecewise linear payoff function will be considered, and the optimal point estimate will be shown to decrease as risk aversion increases. A piecewise linear payoff function with $\Theta \subset A \subset R = (-\infty, \infty)$ is given by

$$g(\theta,a) = \begin{cases} \theta(k_u+k_o) - ak_o & \text{for } \theta \leq a \\ k_u a & \text{for } \theta \geq a. \end{cases} \tag{3.1}$$

with $k_u > 0$ and $k_o > 0$. With a linear utility function U the loss function $L(\theta,a)$ is given by

$$L(\theta,a) = \begin{cases} k_o(a-\theta) & \text{for } \theta \leq a \\ k_u(\theta-a) & \text{for } \theta \geq a \end{cases} \tag{3.2}$$

The interpretation of $L(\theta,a)$ is that the per unit loss due to underestimation is k_u and the loss due to overestimation is k_o. The optimal point estimate $\hat{a}$ is any $k_u/(k_u+k_o)$ fractile of the distribution function of θ (see Raiffa and Schlaifer [7,pp.176-207] and LaValle [5,pp.508-535]). For a nonlinear utility function the loss function is

$$L(\theta,a) = \begin{cases} U(k_u\theta) - U(\theta(k_u+k_o)-ak_o) & \text{for } \theta \leq a \\ U(k_u\theta) - U(k_u a) & \text{for } \theta \geq a \end{cases} \tag{3.3}$$

To demonstrate that the optimal point estimate for the piecewise linear payoff function is decreasing in absolute risk aversion, a result will be presented for a more general payoff function that includes (3.1) as a special case, and then an interpretation of the risk aversion affect will be given.

Theorem 3: Let $U_1(y)$ and $U_2(y)$ be two utility functions such that $r_{U_1}(y) \geq r_{U_2}(y) \geq 0$ for all y and $r_{U_1}(y) > r_{U_2}(y)$ for some $y = g(\theta,\hat{a}_1)$ (or on some interval of such y) with positive probability, where $\hat{a}_1$ is optimal for U_1 and $\hat{a}_2$ is optimal for U_2. For $\theta \geq a$, let $g(\theta,a) \equiv g^*(a)$ be constant in θ, concave in a, and right differentiable with $g^{*\prime}(a) > 0$. For $\theta \leq a$, assume that $g(\theta,a)$ is decreasing in a, concave in a, left differentiable with respect to a, and such that $g(a,a) = g^*(a)$. Then $\hat{a}_1 < \hat{a}_2$.

Proof: The necessary condition for the optimal estimate $\hat{a}_1$ for (3.3) for U_1 may be written as

$$-g^{*\prime}(\hat{a}_1)(1-P(\hat{a}_1)) - \int_{-\infty}^{\hat{a}_1} \frac{U_1'(g(\theta,\hat{a}_1))g'(\theta,\hat{a}_1)}{U_1'(g^*(\hat{a}_1))}\, dP(\theta) = 0. \qquad (3.4)$$

The estimate $\hat{a}_1$ satisfying (3.4) is a unique global optimum for a strictly concave utility function, since $g(\theta,a)$ is concave, $U_1(g(\theta,a))$ is strictly concave in a, and consequently, $EL(\theta,a)$ is strictly convex in (a). For $\theta \in (-\infty,\hat{a}_1)$, $g^*(\hat{a}_1) > g(\theta,\hat{a}_1)$ so letting $y^* = g(\theta,\hat{a}_1)$, substituting $U_2'(g(\theta,\hat{a}_1))/U_2'(g^*(\hat{a}_1))$ for the similar term for U_1 in (3.4), and using (2.7) implies that

$$-g^{*\prime}(\hat{a}_1)(1-P(\hat{a}_1)) - \int_{-\infty}^{\hat{a}_1} \frac{U'_2(g(\theta,\hat{a}_1))g'(\theta,\hat{a}_1)}{U_2'(g^*(\hat{a}_1))}\, dP(\theta) < 0.$$

Since $EL(\theta,a)$ is strictly convex, $\hat{a}_2 > \hat{a}_1$. //

The following Corollary extends Theorem 3 to a parameterization of U.

Corollary: The optimal point estimate is an increasing (decreasing) (constant) function of y, where utility is $U(y+g(\theta,a))$, for U exhibiting decreasing (increasing) (constant) absolute risk aversion.

The piecewise linear payoff function in (3.1) satisfies the hypotheses of Theorem 3 and its Corollary, so the optimal point estimate is decreasing in absolute risk aversion. The interpretation of Theorem 3 is that greater risk aversion implies that losses incurred when $\theta < a$ are given more weight than losses incurred when $\theta > a$, since for U_1 strictly concave marginal utility $U_1'(g(\theta,\hat{a}_1))$ in (3.4) increases as $g(\theta,\hat{a}_1)$ decreases while for $\theta > a$ marginal utility is constant and is less than $U_1'(g(\theta,\hat{a}_1))$. For example, if U_2 is a linear utility function the necessary condition for the optimal estimate $\hat{a}_2$ is

$$-g^{*\prime}(\hat{a}_2)(1-P(\hat{a}_2)) - \int_{-\infty}^{\hat{a}_2} g'(\theta,\hat{a}_2)dP(\theta) = 0.$$

With a strictly concave utility function $g'(\theta,\hat{a}_2)$ is given greater weight for $\theta \in(-\infty,\hat{a}_2)$ than is $g^{*\prime}(\hat{a}_2)$, so evaluated at $\hat{a}_2$ the expression on the left-hand side of (3.4) is positive for U_1 strictly concave. The optimal point estimate for U_1 is thus less than the optimal point estimate for a linear utility function. In terms of the piecewise linear loss function derived from (3.1) greater risk aversion results in an optimal point estimate with a greater probability of underestimation. The decision maker prefers to accept the greater probability of a loss due to underestimation in order to reduce the risk of a loss due to overestimation.

The effect of risk aversion on the optimal point estimate for the piecewise linear payoff function may also be illustrated by examining the form of the loss function in (3.3). The loss function, for a strictly concave utility function with convex marginal utility may be shown to be convex in θ for $\theta < a$ and concave for $\theta > a$. For $\theta > a$ the loss function is concave in θ for a fixed a, since $\frac{d^2L(\theta,a)}{d\theta^2} = k_u U''(k_u\theta) < 0$. For $\theta < a$, $\frac{d^2L(\theta,a)}{d\theta^2} = k_u^2 U''(k_u\theta) - (k_u+k_o)^2$ $U''(\theta(k_u+k_o)-ak_o)$. Since $k_u\theta > \theta(k_u+k_o) - ak_o$ for $\theta < a$, and for U' convex $U''(k_u\theta) \geq U''(\theta(k_u+k_o)-ak_o)$. Then since $(k_u+k_o)^2 > k_u^2$ and U is concave, $\frac{d^2L(\theta,a)}{d\theta^2} > 0$, and the loss function is convex for $\theta < a$.

Pratt indicates that for U nonincreasingly absolute risk averse, $U'U''' \geq (U'')^2$, and hence $U''' \geq 0$ for $U' > 0$, so U' is convex. For U exhibiting increasing absolute risk aversion U' may or may not be convex. Risk aversion thus results in an asymmetric loss function even if the loss function for a linear utility function in (3.2) is symmetric. Furthermore, it is inappropriate to attempt to represent greater risk aversion by changing, for example, from a piecewise linear to a squared-error loss function. One possible rationale for this might be that losses are more serious the farther θ is from a. However, risk aversion implies that differences $\Delta = |\theta - a|$ are weighted more heavily (by the appropriate marginal utility) for $\theta < a$ than for $\theta > a$.

For a linear utility function the optimal point estimate for the piecewise linear payoff function is a constant function of the ratio $\frac{k_u}{k_u+k_o}$, is decreasing in k_o, and is increasing in k_u. With a strictly concave utility function the following results obtain.

Theorem 4: A) For the piecewise linear payoff function in (3.1), $\frac{d\hat{a}}{dk_o} < 0$. B) For the payoff function $kg(\theta,a)$ with $k > 0$ and $g(\theta,a)$ as in Theorem 2, $\frac{d\hat{a}}{dk} = 0$ if U exhibits constant relative risk aversion.

Proof: A) Implicit differentiation of the necessary optimality condition in (3.4) with respect to k_o yields $\frac{d\hat{a}}{dk_o} = -(1/EL_{aa})$ $[\int_{-\infty}^{\hat{a}} (U'(v)+k_o U''(v)(\theta-\hat{a}))dP(\theta)]$, where $v = \theta(k_u+k_o) - \hat{a}k_o$ and EL_{aa} is the (positive) second derivative of the expected loss function evaluated at $\hat{a}$. For $\theta \in (-\infty,\hat{a})$, $(\theta-\hat{a}) < 0$, so $\frac{d\hat{a}}{dk_o} < 0$.

B) Implicit differentiation of the necessary optimality condition with respect to k yields

$$\frac{d\hat{a}}{dk} = (1/EL_{aa})(1/k) \Big[\int_{-\infty}^{\hat{a}} U''(kg(\theta,\hat{a}))kg'(\theta,\hat{a})\, g(\theta,\hat{a})dP(\theta) + g^*{}'(\hat{a})U''(kg^*(\hat{a}))kg^*(\hat{a})(1-P(\hat{a}))\Big]. \tag{3.5}$$

Add the first-order condition from (3.5) inside the [ ] to obtain

$$\frac{d\hat{a}}{dk} = (1/EL_{aa})(1/k) \Big[\int_{-\infty}^{\hat{a}} U'(kg(\theta,\hat{a}))\, \{1-r_U^*(kg(\theta,\hat{a}))\}\, g'(\theta,\hat{a})dP(\theta) + U'(kg^*(\hat{a}))g^*{}'(\hat{a})\, \{1-r_U^*(kg^*(\hat{a}))\}\, (1-P(\hat{a}))\Big].$$

For U(y) constant relative risk averse $r_U^*(kg^*(\hat{a})) = r_U^*(kg(\theta,\hat{a})) = \beta > 0$, and $\frac{d\hat{a}}{dk} = -(1/EL_{aa})((1-\beta)/k) \left.\frac{dEL}{da}\right|_{a=\hat{a}} = 0$, where $\left.\frac{dEL}{da}\right|_{a=\hat{a}}$ is the necessary optimality condition evaluated at $\hat{a}$. //

There is no counterpart of Theorem 4A) for the effect of increases in k_u, because k_u is involved in the payoff function for both $\theta \leq a$ and $\theta \geq a$, and hence the sign of $\frac{d\hat{a}}{dk_u}$ given below is unclear. Letting $z = k_u\hat{a}$, the derivative is

$$\frac{d\hat{a}}{dk_u} = -\,(1/EL_{aa})\,[k_o \int_{-\infty}^{\hat{a}} U''(v)\theta dP(\theta) - (1-P(\hat{a}))U'(z)(1-k_u\hat{a}r_U(z))].$$

For $\Theta = A \subset (0,\infty)$ and U constant relative risk averse, $k_u\hat{a}r_U(z) = r_U^*(z) = \beta$. If $\beta \leq 1$, $\frac{d\hat{a}}{dk_u} > 0$, but for $\beta > 1$, $\frac{d\hat{a}}{dk_u}$ may be negative. Consequently, while an increase in the loss (in payoff) due to overestimation decreases the optimal point estimate, an increase in the loss (in payoff) due to underestimation may increase, decrease, or leave unchanged the optimal estimate. A proportional increase in both k_o and k_u will have no affect on the optimal point estimate if the utility function exhibits constant relative risk aversion.

4. Numerical Examples

This section presents numerical results that indicate the magnitude of the risk aversion effects for the quadratic payoff function. Table 1 presents numerical results for the quadratic payoff function, normal distribution, and exponential utility function $U(y) = -\exp(-\alpha y)$, $\alpha > 0$, example of Section 2. The optimal estimate for U linear is $\hat{a} = \mu = 5$, and the optimal point estimates are decreasing in risk aversion $(r_U(y) = \alpha)$, since $\hat{a} > 0$. The point estimate also is decreasing in the standard deviation, since the risk premium is increasing in σ. Table 2 presents numerical results for the quadratic payoff function, constant relative risk averse utility function $U(y) = y^c = (150 + g(\theta,a))^c$, and normal, uniform, and exponential distributions. To indicate the "degree of risk aversion" of the utility functions, Table 3 presents the certainty equivalent for a lottery that pays \$100 or \$0 each with probability one-half. For example, with $\alpha = .01$ the decision maker is indifferent between \$39.99 with probability one and the lottery. The values of α used in Table 1 reflect very risk averse preferences.

5. Conclusions

Whenever payoffs to the decision maker are involved, estimation must take into account not only the payoffs but also the risk preferences of the decision maker. Considerable evidence has been accumulated indicating that many decision makers are risk averse, and this paper has investigated the effects of risk aversion on optimal point estimates. For the piecewise linear payoff function the optimal point estimate is decreasing in absolute risk aversion, while increased absolute risk aversion moves the optimal point estimate towards the origin for a quadratic payoff function. The magnitude of the risk aversion effect clearly depends on the "degree of risk aversion," and the numerical results of Section 4 indicate that the risk aversion effect can be significant.

Table 1

Optimal Estimates for Quadratic Payoff Function

$r_U(y) = \alpha$; Normal Distribution, $\mu = 5$; $k = .5$

σ	$\alpha=.01$	$\alpha=.05$	$\alpha=.1$	$\alpha=.3$
.5	4.975	4.878	4.762	4.348
1.0	4.951	4.762	4.545	3.846
1.5	4.926	4.651	4.348	3.448
2.0	4.902	4.545	4.167	3.125
3.0	4.854	4.348	3.846	2.632
4.0	4.808	4.167	3.571	2.273
5.0	4.762	4.000	3.333	2.000

Table 2

Optimal Estimates for Quadratic Payoff Function

$$r_U(y) = \frac{1-c}{y} \; ; \; k = .5$$

Normal Distribution, $\mu = 5$

σ	c=.05	c=.1	c=.15	c=.20	c=.4	c=.6	c=.8
1	4.957	4.960	4.962	4.965	4.973	4.982	4.991
3	4.746	4.759	4.772	4.785	4.837	4.890	4.945
5	4.397	4.425	4.454	4.483	4.603	4.728	4.861

Uniform Distribution on [0,10]

4.766	4.778	4.790	4.802	4.850	4.899	4.949

Exponential Distribution, mean 5

4.551	4.572	4.594	4.616	4.708	4.807	4.912

Table 3

Certainty Equivalents for Lottery

$$x = \begin{cases} 0 & \text{with probability } .5 \\ 100 & \text{with probability } .5 \end{cases}$$

Utility Function	$\alpha=.01$	$\alpha=.05$	$\alpha=.1$	$\alpha=.3$
$U(y)=-\exp(-\alpha y)$ $=-\exp(-\alpha(100+x))$	39.99	13.73	6.93	2.31

	c=.05	c=.10	c=.15	c=.2	c.=4	c=.6	c=.8
$U(y)=y^c=(100+x)^c$	41.85	42.27	42.70	43.13	44.85	46.57	48.29

Footnotes

1. Corresponding to every $U(g(\theta,a))$ is a loss function as in (1.2), but a given loss function may correspond to more than one $U(g(\theta,a))$.

2. For $a \in A$ $L(\cdot,a)$, and hence $U(g(\cdot,a))$, is assumed to be P-integrable. For example, if $U(y) = \log(y)$, the probability that y is non-positive is assumed to be zero. If $g(a,\theta)$ may take on negative values, the utility function $U(g(a,\theta))$ is considered to be $\log(y+g(a,\theta))$ where $y > -\min_{\substack{a \in A \\ \theta \in \Theta}} g(a,\theta)$.

3. Arrow [1] has argued that decreasing absolute risk aversion is a reasonable property of utility functions, since that property implies that increased wealth will cause an individual to pay less for insurance against a given risk.

4. The relative risk aversion index is the elasticity of marginal utility.

5. A more general study of price taking firms is provided by Baron [2].

6. In a decision context the expected utility and not the expected loss in payoff is of interest. The latter is considered here in order to relate the results to the usual inference procedures.

7. If $U(z) = -\infty$ for $z \leq 0$, the result holds for $k(y+g(\theta,a))$ where $y+g(\theta,a) > 0$ for all $\theta \in \Theta$ and $a \in A$.

8. The proof follows that given by Arrow [1,pp.119-120].