Loss Aversion and Reference Points in Contracts

D.R. Just

Steve Wu

Paper prepared for presentation at the
SCC-76 Meeting, March 31-April 2, 2005,
Myrtle Beach, South Carolina

Loss Aversion and Reference Points in Contracts

Abstract

Loss aversion has become the dominant alternative to expected utility theory for modeling choice under uncertainty. The setting of the base payment in contracts provides an interesting application of referenced based decision theory. The impact of loss aversion on contract structure depends critically on whether reservation opportunities (outside options) are evaluated with respect to the reference point implied in the contract. We show that when reservation opportunities are independent of the reference point, reward contracts are optimal. However, when reservation opportunities are evaluated against the reference point, then penalty contracts are more efficient.

March, 2005

JEL Codes: L14, D81, D21, D82

In the standard contracting problem with hidden actions, a principal hires an agent to perform a task which affects the principal's revenues. Because revenue is stochastically related to the amount of effort exerted by the agent, and effort is costly to the agent, the optimal contract requires that the agent receives payments that are contingent on performance. The specific structure of the pay for performance scheme, which can affect marginal incentives for effort, is dependent on the agent's risk preference. A well known result is that when agents are highly risk averse, the optimal contract involves making pay less dependent on performance as the agent's cost of risk bearing will be high making it expensive for the principal to motivate the agent using variable pay. When the agent is relatively risk tolerant, payment to the agent should be more variable to provide high powered incentives.[1]

While contract theorists have focused much of their attention on the proper structure of marginal incentives for effort, it may be important to understand how a principal should determine the base pay, as well as when to use penalties versus rewards in structuring incentive schemes. In practice, incentive contracts consist typically of a base level of pay, and some schedule of rewards and penalties based upon performance objectives. For example, Hueth and Ligon (2003) examine numerous supply contracts from the processing tomato industry and find that these contracts typically provide a base price for each ton of tomatoes, along with various penalties and bonuses that are contingent on various quality measures. Horstmann, Mathewson, and Quigley (2002) report that contracts for life insurance sales agents contain bonuses for policy sales and renewals, and penalties for policy lapses. Bonuses and penalties are also frequently

[1] A similar result is obtained when the agent is risk neutral but faces a limited liability constraint (Innes, 1990). The main difference is that, instead of having to compensate the agent for risk bearing, the optimal contract would pay the agent limited liability rents in order to provide incentives.

observed in general sales agent commission contracts (Tallitsch and Moynihan, 1994). While these examples illustrate that pay for performance is clearly used in practice, just as standard agency theory predicts, the theory cannot explain why these contracts use rewards rather than penalties (or vice versa) and how base prices are determined.

There is good reason for ignoring such topics. Standard contract theory employs expected utility theory, which assumes that there is a well defined level of utility associated with each possible level of wealth. In this context, it matters little whether a high base pay is combined with penalties or a low base pay is combined with rewards so long as the amount of total pay is the same for each level of performance (Lazear 1998). For example, a piece rate scheme paying 5 cents per piece should, according to standard theory, produce the same effort level and profit levels as paying $20 for 400 minus a penalty of 5 cents for each piece under 400. In this example, the second contract simply draws attention to a specific performance level, without changing the incentive structure. Because expected utility cannot assign different levels of utility to a single level of wealth, it cannot differentiate between these two contract schemes. In short, conventional theory suggests that bonuses and penalties are perfect substitutes.

While conventional theory offers us no insights into how contract designers should establish the base pay, performance levels, and penalties and rewards, behavioral economic research might offer us clues. For example, Kahneman and Tversky (1979) suggest that reference points can play a significant role in affecting people's utility. Kahneman and Tversky assert that people not only care about their absolute wealth level, but are also concerned about how their wealth deviates from some reference level of wealth and may be more averse to losses (relative to the reference point) than gains of the same size. Under these assumptions, the setting of the base price can provide a reference point, while rewards and penalties represent deviations from the reference point. In this case, penalties and rewards are no longer perfectly substitutable and the setting of the base price can interact with marginal incentives in a non-trivial way.

In this paper we explore the theory of reference points and its implications for the determination of the base pay, and penalties and rewards in simple principal agent problems. Our key assumption is that the agent is *loss averse* (Kahneman and Tversky, 1979) rather than merely risk averse, so that an increase in wealth has less impact on utility than does an equal decrease in wealth, where increases and decreases are measured against some reference level of wealth. This reference level of wealth is largely determined by the base pay in the contract so that losses and rewards are measured relative to the base pay. Because the value function of the loss averse agent is typically concave above the reference point and convex below it, the careful design of the correct base pay can affect the marginal utility of wealth and will therefore have a non-trivial impact on risk (loss) premiums and incentives for effort.

While choosing the right base pay can pin down the reference point and impact incentives under loss aversion, it is well known amongst prospect theorists that reference points are sensitive to context and framing. We therefore examine two hypothetical cases that might be of relevance in contracting.

In the first case, we assume that the agent evaluates all outcomes, including reservation opportunities (outside options), against the reference point (base pay) given in the contract. If the agent makes such a comparison, then the optimal contract would be characterized by a high base price, along with penalties for poor performance. The intuition for this result is that, because the utility function is steeper over losses, the *same* pay for performance schedule can provide greater marginal incentives if it is designed to operate over the loss portion rather than the gain domain of the utility function. Moreover, a high base pay combined with punishments will push mean compensation below the base pay where the agent will be risk loving rather than risk averse. Because the agent tends to be risk loving over losses, the standard tradeoff between risk and incentives is no longer true; instead, risk now *complements* incentives, enabling the principal to provide strong incentives even if the relationship between effort and performance is noisy.

In the second case, we assume *reservation independence* so that the agent does not evaluate outside options against the reference point. Instead, the reference point is used to evaluate outcomes *within context* so that reservation opportunities fall outside of the context of the contract. In this case, outside opportunities are measured objectively (gains and losses are not exaggerated) and are not compared to the reference point in the contract. In this scenario, the optimal contract provides the agent with a relatively low base pay combined with rewards, on average. The intuition is that, because both rewards and losses *within* the context of the contract are exaggerated, rewards produce a utility level that exceeds the "objective" value (recall reservation utility is objectively assessed), whereas penalties would produce a utility level that is below the objective level. Given this exaggeration, it is cheaper to provide incentives via rewards, as it would be very costly for the principal to use penalties to motivate effort while having to ensure that the contract yields ex ante utility (which is exaggerated downward) that exceeds the reservation utility (which is not exaggerated).

Contracting With Moral Hazard

The standard principal agent model with moral hazard has become a workhorse model for describing many economic relations in insurance, labor markets, CEO compensation, organizational theory, sharecropping, and various other business related fields.

A general formulation of the model (e.g. Holmstrom 1979) involves a principal who contracts with an agent to perform some task which affects the output (or revenue), $q \in [\underset{\smile}{q}, \hat{q}]$, desired by principal. The agent produces output by exerting some non-observable and non-verifiable effort, $e \in \xi \subseteq \square$, which is stochastically related to output via the cumulative distribution function $H(q \mid e)$, which has a conditional density $h(q \mid e)$. Exerting effort is costly for the agent in terms of disutility. The effort cost function is given by $z(e)$, which satisfies the conditions $z'(e) > 0,\ z''(e) \geq 0$ with $z(0) = 0$. The principal thus faces the problem

(1) $$\max_{e,\{w(\tilde{q})\}} \int V\left(\tilde{q} - w\left(\tilde{q}\right)\right) h\left(\tilde{q} \mid e\right) d\tilde{q},$$

subject to

(2) $$\int U\left(w\left(\tilde{q}\right)\right) h\left(\tilde{q} \mid e\right) d\tilde{q} - z\left(e\right) \geq \underline{U},$$

and

(3) $$e \in \arg\max_{\tilde{e}} \int U\left(w\left(\tilde{q}\right)\right) h_e\left(\tilde{q} \mid e\right) d\tilde{q} - z_e\left(e\right),$$

where $V\left(\cdot\right)$ is the principal's utility function, and $w(\tilde{q})$ represents the contractually specified transfer from the principal to the agent, and $U\left(\cdot\right)$ is the agent's utility function. Rogerson (1985) has shown that if $H(q \mid e)$ satisfies the monotone likelihood ratio property (MLRP) and the convexity of the distribution function condition (CDFC), then the constraint (3) can be replaced by the agent's first order condition:

(4) $$\int U\left(w(\tilde{q})\right)h_e\left(\tilde{q}\,|\,e\right)d\tilde{q}-z_e\left(e\right)=0$$

The solution to the contract design problem is well known and is characterized by the following equation:

(5) $$\frac{V'\left(\tilde{q}-w(\tilde{q})\right)}{U'\left(w(\tilde{q})\right)}=\theta+\mu\frac{h_e\left(\tilde{q}\,|\,e\right)}{h\left(\tilde{q}\,|\,e\right)},$$

where μ and θ are non-negative multipliers. This equation essentially describes the relationship between pay and performance. Under MRLP, $\frac{h_e\left(\tilde{q}\,|\,e\right)}{h\left(\tilde{q}\,|\,e\right)}$ is increasing in $\tilde{q}$ so that a high output sends a signal to the principal that the agent has exerted high effort. In this case, the agent is rewarded with a high payment. On the other hand, when output is low, the agent receives a low transfer. The degree to which transfers vary with output will also depend on the relative curvatures of the utility functions of the respective parties. For example, if the agent is extremely risk averse relative to the principal, then transfers will be less sensitive to output variation because it would be more efficient for the principal to bear more of the risk by reducing the variation in transfers.

From (5) we may derive many of the properties of the wage/quality relationship. While this contract is useful in discussing the risk sharing effects of contracts, and provides a rationale for pay for performance when moral hazard is present, it offers little guidance on practical matters such as how a base payment is to be determined, and whether premiums or deductions should be used in the pay for performance scheme. For example, most contracts specify some base pay, several deductions for poor performance, and several premia for good performance. In the context above, the contract may look like

(6) $$w(\tilde{q}) = \underset{\smile}{w} + f\left(\tilde{q} - \underset{\smile}{q}\right),$$

where f is a non-negative valued function, $\underset{\smile}{w}$ is the lowest possible wage, and the support of $\tilde{q}$ is given by $\left[\underset{\smile}{q}, \widehat{q}\right]$. The standard theory suggests no specific reason to use a premium rather than a deduction, as the same utilities and payoffs can be achieved using either. For example, if the above contract were optimal and represented a premium paid for good performance, then we could define $\widehat{w} = \underset{\smile}{w} + f\left(\widehat{q} - \underset{\smile}{q}\right)$, and the contract

(7) $$w(\tilde{q}) = \widehat{w} - \left[\widehat{w} - \left(\underset{\smile}{w} + f\left(\tilde{q} - \underset{\smile}{q}\right)\right)\right] = \widehat{w} - g\left(\widehat{q} - \tilde{q}\right),$$

would yield the same payoff to both parties in all cases.

Despite the fact that theory suggests no particular reason for the use of premia or deductions, there appears to be significant thought and gaming involved in choosing the base level of pay in real world contracts. Hueth and Ligon (2002) note that, while some processing tomato contracts contain special premia written in processor specific contracts, the boilerplate contracts used for all tomato farmers contain only deductions. Curtis and McCluskey (2003) find a mix of premia and deductions used in production of processing potatoes. Thus, there may be some behavioral phenomena that impacts the structure of these contracts. Several behavioral models may lead to the use of specific base payments. The model that has had the greatest impact on the profession, and has been applied most ubiquitously, seems a good starting point for our analysis. We propose that loss aversion on the part of the agent may drive the structure of premia and deductions, allowing the principal to obtain greater effort through manipulating the agent's reference payoff.

Loss aversion has become the preferred behavioral model to describe behavior dealing with risk. Kahneman and Tversky first proposed loss aversion as part of their prospect theory (1979). Loss aversion supposes that individuals experience diminishing marginal utility of gains in wealth, but also diminishing marginal pain from losses. Thus, a utility of wealth function must be contingent on a reference point, against which gains and losses are measured. Above this reference point, the utility function is concave, reflecting risk averse behavior. Below this reference point, individuals behave as if risk loving, willing to risk lower returns for a chance at returning to their reference point.

There are many reasons why a principal may desire to manipulate the reference level of wealth for an agent. First, by doing so, he may manipulate the marginal utility of income, thus making his marginal incentive more effective. Secondly, Sandmo's classic result (1971) suggests that risk attitude can affect input efficiency, and expected profit. Thus the principal may be able to enhance profits by manipulating the risk attitude of the agent via the reference payout level.

We can write the loss averse value function as

(8) $$u(x \mid w) = \begin{cases} v^{+}(x-w) & \text{if} \quad x > w \\ v^{-}(x-w) & \text{if} \quad x \le w \end{cases}$$

where $v^{+}(0) = v^{-}(0) = 0$, so that utility is continuous, and $v^{+}{}''(s) < 0, v^{-}{}''(s) > 0$. Figure 1 displays an example of what the referenced based utility function may look like. Note that the function is not differentiable at the reference point, and declines steeply when moving into the loss domain. The loss aversion paradigm has found support in many contexts, including experimental (Tversky and Kahneman, 1992; Camerer, 1995) and non-experimental (Benartzi and Thaler, 1995) contexts, and even in contexts that do not involve risk (Kahneman Knetsch and Thaler, 1991).

A common criticism of prospect theory is the problem of determining the reference point. In contracting, the base pay serves as a natural reference point. However, because context and framing are also important in prospect theory, an important consideration in contract design is how agents view reservation activities (outside options). In standard contract theory, the reservation utility is typically treated as fixed and merely serves as a constraint on the principal's contract design problem. In prospect theory, the outside options may play a more important role depending on whether an agent evaluates these outside options against the reference point in the contract, or not. We will illustrate the importance of this point in subsequent discussions by examining two cases - one where the agent compares outside options to the reference point and a second case where the agent does not make this comparison so that the reservation utility is independent of the reference point.

Contracts and Reference Points

If loss aversion is important in risky behavior, then principals should have a strong interest in manipulating reference points. In order to illustrate this principle, we propose the following model of agent behavior, based on the prospect theory value function

$$\int U\left(w \mid \bar{w}\right) h\left(\tilde{q} \mid e\right) d\tilde{q} - e, \tag{9}$$

where, as before, U is a measure of utility of wealth, w, given a level of base pay, $\bar{w}$. The disutility of effort is now assumed to be linear, which can be made without loss of generality. We assume that $U_{ww}\left(w \mid \bar{w}\right) < 0$ for $w > \bar{w}$, $U_{ww}\left(w \mid \bar{w}\right) > 0$ for $w < \bar{w}$.

Further, $\lim_{w \uparrow \bar{w}} U(w \mid \bar{w}) = \lim_{w \downarrow \bar{w}} U(w \mid \bar{w}) = 0$, and $\lim_{w \uparrow \bar{w}} U_w(w \mid \bar{w}) > \lim_{w \downarrow \bar{w}} U_w(w \mid \bar{w})$, thus the function is continuous, but not differentiable at the reference level of wealth. Lastly, because we have assumed MRLP of $h(\tilde{q} \mid e)$, this implies that $\frac{\partial E(\tilde{q} \mid e)}{\partial e} > 0$, or that increasing effort increases output on average.

We will also assume that the principal is risk neutral and behaves rationally (in accordance with expected utility theory). The assumption of risk neutrality of the principal is frequently made in the literature and is justifiable if the principal is able to diversify its risks away (e.g. shareholders), or is a larger company that has resources to both diversify its operations and conduct sophisticated market analysis for decision support. The agent, on the other hand, may represent a worker, a small supplier, a farmer or some other entity that has limited resources to diversify or access sophisticated decision support knowledge. Thus, the agent's behavior may be more heavily influenced by risk attitudes and behavioral anomalies.

Following the contract theory literature, we suppose that the principal has all the ex ante bargaining power, and designs the contract to ensure that the agent obtains an expected payoff that, at minimum, covers his reservation utility. However, because context and framing are important notions in the loss aversion literature, it matters whether the agent measures these reservation opportunities against the reference point in the contract or not. In the next two subsections we outline the implications of comparing the reservation opportunity to the reference point (reservation in reference) or examining the reservation opportunity independent of the reference point (reservation independence).

Reservation in Reference

If the agent evaluates his reservation activities in comparison to the reference point given in the contract, the principal must solve the following problem in designing the contract

(10) $$\max_{e^*,\{w(\tilde{q})\},\overline{w}} \int \left(\tilde{q} - w(\tilde{q})\right) h\left(\tilde{q} \mid e^*\right) d\tilde{q},$$

subject to

(11) $$e^* \in \arg\max_{e} \int U\left(w(\tilde{q}) \mid \overline{w}\right) h\left(\tilde{q} \mid e\right) d\tilde{q} - e,$$

(12) $$\int U\left(w(\tilde{q}) \mid \overline{w}\right) h\left(\tilde{q} \mid e^*\right) d\tilde{q} - e^* \geq U\left(w_r \mid \overline{w}\right).$$

Here w_r represents some reservation payoff that generates the reservation utility level. Also, note that the principal now must choose the optimal effort level, $e*$, the optimal contract, $w\left(\tilde{q}\left(e^*\right)\right)$, as well as the base pay, $\overline{w} = w(\overline{q})$, which serves as the reference point. Because this is a particularly difficult problem to solve, we will first examine the optimal contract in the absence of uncertainty. In this case, the above model reduces to

(13) $$\max_{e^*,\{w(\tilde{q})\},\overline{w}} \left(\tilde{q}\left(e^*\right) - w\left(\tilde{q}\left(e^*\right)\right)\right),$$

subject to

(14) $$e^* \in \arg\max_{e} U\left(w\left(\tilde{q}(e)\right) \mid \overline{w}\right) - e,$$

(15) $$U\left(w\left(\tilde{q}(e)\right) \mid \overline{w}\right) - e^* \geq U\left(w_r \mid \overline{w}\right).$$

Given that the agent has signed the contract, she will optimize by exerting effort e, where $U_w\left(w \mid \overline{w}\right) = \frac{1}{w_q \tilde{q}_e}$, if $q(e) \neq \overline{q}$. This leads us to our first proposition.

Proposition 1. Let U be a prospect theoretic value function and suppose that for any Δ we can find $w_l < \overline{w}$ such that $U(w_l \mid \overline{w}) - U(w_l - \Delta \mid \overline{w}) > U(w \mid \overline{w}) - U(w - \Delta \mid \overline{w})$ for any $w > \overline{w}$. Then under the optimal contract without uncertainty, $w(\tilde{q}(e^*)) = w(\overline{q})$.

Proof: Suppose that under the optimal contract $w(\tilde{q}(e^*)) \geq w(\overline{q})$, with $U_w(w \mid w(\overline{q})) = \frac{1}{w_q \tilde{q}_e}$. Consider an alternative reference wealth $\overline{q}_a = \tilde{q}_e + \varepsilon$, $\varepsilon > 0$.

Because, $\lim_{w \uparrow \overline{w}} U_w(w \mid \overline{w}) > \lim_{w \downarrow \overline{w}} U_w(w \mid \overline{w})$, the maximum of the first derivative occurs as one approaches the reference point from the left. Because $\lim_{w \uparrow \overline{w}} U_w(w \mid w(\overline{q}_a)) > \frac{1}{w_q \tilde{q}_e}$, and $\lim_{w \downarrow \overline{w}} U_w(w \mid w(\overline{q}_a)) > \frac{1}{w_q \tilde{q}_e}$, the optimal level of effort for the agent must be larger under the new reference level of wealth, if ε is small enough, and the optimal level of effort is not 0. This latter possibility is excluded if

$U(w(\tilde{q}(e^*)) \mid w(\overline{q}_a)) - U(w_r \mid w(\overline{q}_a)) > U(w(\tilde{q}(e^*)) \mid \overline{w}) - U(w_r \mid \overline{w})$. ■

To understand proposition 1, note that the principal essentially must determine some optimal effort level *e** which can be implemented with a contract payment of $w(\tilde{q}(e^*))$. The principal must also determine some optimal reference quality level, $\overline{q}$, which will allow the principal to set some optimal base pay, $\overline{w}$. Proposition 1 essentially tells us that if the agent is loss averse, then the principal's choice of *e** also results in the optimal choice of the reference quality, $\overline{q}$, and vice versa. The intuition is that effort is monotonically increasing with the slope of the utility function. The slope is increasing to its maximum as the reference point is approached from the left. The slope is decreasing from something less than its maximum as the reference point is approached from the right.

Proposition 1 also implies that if the loss portion of the value function has less curvature than the gains portion, then the optimal payout will be the base pay. The constraint placed on the convexity of the loss function is a very minimal requirement that is met by all value functions currently used in the literature (see for example Tversky and Kahneman, 1992). This restriction simply requires that the pain from any loss be larger than the pleasure from an equivalent gain. This is of course one of the two main hypotheses behind loss aversion. Thus if agents are loss averse, the principal will optimize by stating the intended effort as resulting in exactly the base pay, this being the point with the greatest marginal utility of wealth.

Returning to the case of uncertainty given in (10) – (12), only complicates the picture mildly, with the result depending on the precision of the quality signal. To see this, note that if the conditional distribution of output satisfies MLRP and CDFC, then we can replace constraint (10) with the constraint:[2]

(16)
$$\int U\left(w(\tilde{q}) \mid \overline{w}\right) h_e\left(\tilde{q} \mid e\right) d\tilde{q} - 1 = 0 .$$

Totally differentiating the agent's first order condition with respect to e and $\overline{w}$ yields

(17)
$$\frac{de}{d\overline{w}} = -\frac{\left[\int U_{\overline{w}}\left(w(\tilde{q}) \mid \overline{w}\right) h_e\left(\tilde{q} \mid e\right) d\tilde{q}\right]}{\left[\int U\left(w(\tilde{q}) \mid \overline{w}\right) h_{ee}\left(\tilde{q} \mid e\right) d\tilde{q}\right]} .$$

[2] It is not difficult to prove that the first order approach is valid even with a loss averse utility function so long as MRLP and CDFC are satisfied.

Here, the expression in the denominator must be negative for the agent's optimization to hold. Thus, maximizing effort with respect to $\bar{w}$ must occur where $\frac{de}{d\bar{w}} = 0$.[3] Note that altering the reference wealth exerts no cost to the principal, yet will result in increased mean profits, so long as the average wage remains constant. Thus, the principal should maximize effort over the reference level of pay.

The denominator is just the agent's second order condition, which must be negative. Thus, the optimal $\bar{w}$ solves

(18) $$\left[\int U_{\bar{w}}\left(w(\tilde{q}) \mid \bar{w}\right) h_e\left(\tilde{q} \mid e\right) d\tilde{q}\right]\left[\int U\left(w(\tilde{q}) \mid \bar{w}\right) h\left(\tilde{q} \mid e\right) d\tilde{q} - e^* - U\left(w_r \mid \bar{w}\right)\right] = 0,$$

the corresponding complementary slackness condition (because $\bar{w}$ does not enter directly into the expected profit of the principal). The expression $U_{\bar{w}}\left(w(\tilde{q}) \mid \bar{w}\right)$ is always negative. Raising the reference point by some amount Δ is exactly equivalent to lowering the wage by Δ; thus raising the reference point always lowers the value of any gamble.

Proposition 2 Let U be a prospect theoretic value function and suppose that for any Δ we can find $w_l < \bar{w}$ such that $U\left(w_l \mid \bar{w}\right) - U\left(w_l - \Delta \mid \bar{w}\right) > U\left(w \mid \bar{w}\right) - U\left(w - \Delta \mid \bar{w}\right)$ for any $w > \bar{w}$, and that the conditional distribution of output be of the form $h\left(\tilde{q}(e) \mid e\right) = h\left(\tilde{q}(e) - g(e)\right)$, satisfying MLRP, and CDFC, with $g'(\cdot) > 0$, $g''(\cdot) < 0$. Then under the optimal contract, $E\left(w \mid e^*\right) \le \bar{w}$ if $U''\left(w \mid \bar{w}\right) < k$ for some k and for any $w < \bar{w}$.

[3] Note that we assume that this function is concave. If it were not so, then effort can be raised to infinity simply by changing the reference point (no change in pay scale would be needed).

Proof: First, the MLRP and CDFC constraints require that $h'(\cdot)>0$. We can show that the IR constraint is relaxed by penalties. To see, this, note that for any w, $w-w_r>0$, the value $U(w\,|\,\overline{w})-U(w_r\,|\,\overline{w})$ is maximized where $\overline{w}=w$. Thus, the constraint is least restrictive when the certainty equivalent is such that $CE=\overline{w}$. Note, $CE>E(w\,|\,e^*)$ if $E(w\,|\,e^*)<\overline{w}$ by convexity, and $CE<E(w\,|\,e^*)$ if $E(w\,|\,e^*)\geq\overline{w}$ by concavity. The value function's change in slope at $\overline{w}$ means that the function will behave as if concave, satisfying Jensen's inequality, if the value function is not too convex over losses. Thus, by the continuous nature of the CE, it can only equal $\overline{w}$ if $E(w\,|\,e^*)<\overline{w}$. Secondly, we can show that for any given wage schedule, effort increases as $\overline{w}$ is increased from $E(w\,|\,e^*)$. Totally differentiating the incentive compatibility constraint, we find

$$\frac{de}{d\overline{w}}=-\frac{\left[\int U_{\overline{w}}\left(w(\tilde{q})\,|\,\overline{w}\right)h_e\left(\tilde{q}\,|\,e\right)d\tilde{q}\right]}{\left[\int U\left(w(\tilde{q})\,|\,\overline{w}\right)h_{ee}\left(\tilde{q}\,|\,e\right)d\tilde{q}\right]}=\frac{\left[\int U_{\overline{w}}\left(w(\tilde{q})\,|\,\overline{w}\right)h'\left(\tilde{q}-g(e)\right)g'(e)\,d\tilde{q}\right]}{\left[\int U\left(w(\tilde{q})\,|\,\overline{w}\right)h_{ee}\left(\tilde{q}\,|\,e\right)d\tilde{q}\right]}>0$$

The denominator is the second order condition and must be negative, the integrand in the numerator must be negative. Thus, the optimal $\overline{w}$ solves

(19) $$\left[\int U_{\overline{w}}\left(w(\tilde{q})\,|\,\overline{w}\right)h_e\left(q\,|\,e\right)dq\right]\left[\int U\left(w\,|\,\overline{w}\right)h\left(q\,|\,e\right)dq-e^*-U\left(w_r\,|\,\overline{w}\right)\right]=0,$$

the corresponding complementary slackness condition (because $\overline{w}$ does not enter directly into the expected profit of the principal). Equation (19) can only be satisfied where the IR constraint binds. Thus, the principal's problem is solved where $CE=\overline{w}$, implying that $E(w\,|\,e^*)<\overline{w}$ ∎

Proposition 2 states the conditions under which we expect the average wage to be below the reference, or base pay, which can only be achieved if significant penalties are in place in the contract. There are two primary reasons that penalties should prevail. First, the utility function is steeper over the loss domain, meaning that a given pay for performance scheme can have a greater impact on marginal effort. Second, as average wage is moved to the left of the reference point, the individual becomes more and more risk loving, increasing the certainty equivalent relative to average wage. This means, that the standard tradeoff between risk and incentives becomes weakened so that the cost of providing incentives to the agent decreases significantly. This gives greater leverage to the principal in providing incentives even if the relationship between output and effort is very noisy.

Figure 2a and 2b illustrate how the principal can reduce average payout below the reservation wage and obtain the same level of effort. This is accomplished by raising the base level of pay, thus shifting the reference point to the right. When the reservation and reference level of pay are equal (Figure 2a) the function behaves as if concave, yielding utility below that obtained from the average wealth. When instead the reference point is shifted up (Figure 2b) the function behaves as if convex above the reservation wage, yielding a higher level of utility for the contract with the same level of expected pay. The irony of this result is that we used a model of loss aversion to show that losses are preferred to gains. As we will show in the subsequent section, this irony derives from how the reservation wage is compared to the base level of pay. Thus, any agent comparing reservation opportunities to the base level of pay, according to prospect theory, will actually behave 'as if' loss loving. The principal can take advantage of the risk loving portion of the utility function by inducing losses relative to the base pay.

Reservation Independence

Here we examine the case where the reference point is only used to evaluate outcomes *within context* – that is, only outcomes that would occur under the contract are compared to the reference point. Outside options that affect the reservation utility are outside the context of the contract and are no longer evaluated against the reference point (base price). Because reservation activities fall outside the context of the contract, they are now measured "objectively" (without exaggeration of gains or losses) with respect to the reference point. Figure 3 depicts the model we propose. Note that the reference based utility is measured in addition to an objective utility function. In other words, outcomes that occur under the contract are mapped into a utility value in accordance with the utility function $U(w\,|\,\overline{w})+u(w)$, whereas outcomes that occur in alternative activities are mapped into a utility value determined by the function $u(w)$. We call $u(w)$ "objective" because it does not magnify outcomes away from the reference point in the same way that $U(w\,|\,\overline{w})+u(w)$ does. Intuitively, an individual feels worse obtaining a $10 base bay and a $5 penalty than they do with an alternative that pays $5 always. The alternative that pays $5 always is not subject to the same framing. There is no sense that the individual did not do as well as they could have, and thus no added costs of loss aversion reflected in the individual's evaluation of this alternative. Figure 3 depicts that $U\left(w\,|\,\overline{w}\right)=0$ for w far enough above $\overline{w}$. This condition is required for a solution to the principal agent problem.

This model may seem contrary to the traditional loss aversion model proposed by Kahneman and Tversky (1979). However, one of the primary principles behind Kahnman and Tversky's argument is that context matters when determining a reference point. One could also suppose that the reservation utility is compared to some latent reference point not specified in the model. Hence, despite its involved nature, this model does reflect the loss aversion phenomenon described in the literature.

If the reservation utility level is independent of the reference wealth, we can rewrite the principal's problem as

(20) $$\max_{e^*,\{w(\tilde{q})\},\overline{q}} \int \left(\tilde{q} - w(\tilde{q})\right) h\left(\tilde{q} \mid e^*\right) d\tilde{q},$$

subject to

(21) $$e^* \in \arg\max_{e} \int \left[u\left(\overline{w}\right) + U\left(w\left(\tilde{q}\right) \mid \overline{w}\right) \right] h\left(\tilde{q} \mid e\right) d\tilde{q} - e,$$

(22) $$\int \left[u\left(\overline{w}\right) + U\left(w\left(\tilde{q}\right) \mid \overline{w}\right) \right] h\left(\tilde{q} \mid e^*\right) d\tilde{q} - e^* \geq \underline{U},$$

where $u\left(\overline{w}\right)$ measures the utility of wealth at the reference level of wealth. Here, it is anticipated that all outcomes that occur under the contract will be evaluated against the base pay whereas all outcomes that occur under an outside option will not be measured against the base pay. Thus the reservation utility is measured without respect to a reference point. This model is consistent with the notion that the individual anticipates that he will behave in a loss averse manner if he accepts the contract. The constraints in (21) and (22) can be rewritten as

(21) $$\int \left[U\left(w\left(\tilde{q}\right) \mid \overline{w}\right) \right] h_e\left(\tilde{q} \mid e\right) d\tilde{q} - 1 = 0$$

(22) $$\int \left[U\left(w\left(\tilde{q}\right) \mid \overline{w}\right) \right] h\left(\tilde{q} \mid e^*\right) d\tilde{q} \geq \underline{U} + e^* - u\left(\overline{w}\right)$$

The certainty equivalent (to the agent) of a contract is given by

(23) $$CE = u^{-1}\left(\int \left[U\left(w\left(\tilde{q}\right) \mid \overline{w}\right) \right] h\left(\tilde{q} \mid e^*\right) d\tilde{q} + u\left(\overline{w}\right) - e^* \right)$$

Differentiating with respect to $\overline{w}$ yields

(24)

$$\frac{dCE}{d\overline{w}} = u_1^{-1}\left(\int\left[U\left(w(\tilde{q})\,|\,\overline{w}\right)\right]h\left(\tilde{q}\,|\,e^*\right)d\tilde{q} + u\left(\overline{w}\right) - e^*\right)$$
$$\times\left(\int\left[-U'\left(w(\tilde{q})\,|\,\overline{w}\right)\right]h\left(\tilde{q}\,|\,e^*\right)d\tilde{q} + u'\left(\overline{w}\right) + \left(\int\left[U\left(w(\tilde{q})\,|\,\overline{w}\right)\right]h_e\left(\tilde{q}\,|\,e^*\right)d\tilde{q} - 1\right)\frac{de}{d\overline{w}}\right)$$
$$= \frac{1}{u'(CE)}\left(\int\left[-U'\left(w(\tilde{q})\,|\,\overline{w}\right)\right]h\left(\tilde{q}\,|\,e^*\right)d\tilde{q} + u'\left(\overline{w}\right)\right)$$

This leads to our next proposition.

Proposition 3 Let U be a prospect theory value function, with

$u'(\overline{w}) < \int_{-\infty}^{\infty}\left[U'\left(w(\tilde{q})\,|\,E\left(w\,|\,e^*\right)\right)\right]h\left(\tilde{q}\,|\,e^*\right)d\tilde{q}$ and $h\left(\tilde{q}(e)\,|\,e\right)$ satisfy MLRP and CDFC. Then

$E\left(w\,|\,e^*\right) \geq \overline{w}$.

Proof: If $\frac{dCE}{d\overline{w}} < 0$, the individual rationality constraint can be relaxed by lowering the reference wealth $\overline{w}$. Let $w(q)$ be the solution to our problem. If the individual rationality constraint is binding on the agent's problem when we set $\overline{w} = E\left(w(q)\,|\,e^*\right)$, then the optimal reference wealth has $E\left(w\,|\,e^*\right) \geq \overline{w}$. This will occur if

$u'(\overline{w}) < \int_{-\infty}^{\infty}\left[U'\left(w(\tilde{q})\,|\,E\left(w\,|\,e^*\right)\right)\right]h\left(\tilde{q}\,|\,e^*\right)d\tilde{q}$. ■

Proposition 3 states that, with reservation independence where only gains and losses within context are exaggerated, the principal should, under most reasonable circumstances, offer rewards on average so that expected payoffs exceed the base price. Figure 4 illustrates that if the reference point is set equal to average pay, the certainty equivalent will be lower than the situation where the reference point is set below mean pay, where w_1 represents a base pay set at or above average pay, and w_2 represents a base pay set below average pay. While both value functions are effectively concave around the mean level of pay, the value function with reference point below the mean (w_2) is always above the value function with reference point at average pay (w_1). Because the individual exaggerates gains as well as losses, it will be cheaper for the principal to increase the agent's within context utility by using rewards rather than punishments. Rewards induce a utility that is above the "objective" reservation utility, whereas punishments would lower the agent's utility relative to the reservation utility. Hence, it becomes much cheaper to induce participation in the contract by using a low base pay combined with rewards rather than a high base pay combined with punishments.

As mentioned previously, for this model to have a finite solution, it must be the case that $u(\bar{w})+U(w\,|\,\bar{w})\le u(w)$ for $w>\hat{w}$, for some $\hat{w}>\bar{w}$. Without this condition, the further to the right of the reference point, the greater the utility. Thus, the principal could set the base pay at negative infinity, and obtain infinite effort. This restriction implies that the principal can only fool the agent to a certain extent, before the agent recognizes that a large bonus is offset by the severely low base pay.

Discussion and Conclusion

While the standard principal agent model sheds light on the shape of the optimal contract, it offers little guidance to contract designers on how to set the base pay, and does not distinguish between punishments and rewards in providing marginal incentives. This paper extends the basic principal agent framework by incorporating behavioral considerations based on prospect theory, resulting in a model that can shed light on why it matters whether penalties are used instead of rewards and vice versa. Under the assumption of loss aversion, the reference point or base payment of the contract can affect the way agents evaluate gains and losses, which can in turn alter both their behavior and the way they evaluate contracts. However, the way reference points affect incentives for effort and participation in the contract will also depend on context and framing. When the agent compares all outcomes, including outcomes that would occur under outside options (options that affect the reservation utility), to the reference point, the principal can maximize profits by offering the agent a relatively high base payment combined with penalties, on average. On the other hand, when only outcomes that occur under the contract are evaluated against the reference point, a low base pay combined with premia should prevail.

Our results suggest that whether rewards or premia should be used in the optimal contract depends partly on the scope of influence of the reference point or base payment in the contract. When opportunities beyond the contract are also assessed against the reference point, there is a strong rationale for the principal to exploit an agent's loss aversion and use penalty contracts. However, it is difficult to imagine that the scope of influence of the reference point in a contract will extend far beyond the contract. It may be more reasonable to assume that the contractually specified reference point has its greatest impact on outcomes under the contract and its sphere of influence will gradually taper off as the agent evaluates opportunities outside the contract. While there is little research that we are aware of to shed light on the scope of influence of reference points, it does appear from casual observation that very few contracts in practice induce average payments that fall below the base salary. For example, many labor contracts include a starting salary and bonuses, and it is rare for average gross pay to fall below the base salary. Managers and CEO's are offered a base salary and then can earn additional bonuses once performance targets are met. In professional sports contracts, it is rare to observe negative incentives where a player has to return part of his base pay to the team if he fails to meet performance objectives. Outside of labor markets, Hueth and Ligon (2003) analyze processing tomato contracts and find that these contracts typically include both premiums and deducts. However, average compensation for a typical contract in 1998 exceeds the base price by $1 per ton. Curtis and McCluskey (2003) examine actual Russet Burbank potato contracts and outcomes for two Washington potato processors. While both deducts and premiums are observed in the contracts, the evidence shows that payoffs consistently exceed the base price for a truckload of potatoes.

Possible avenues for future research include investigating the scope of influence of reference points. For example, if an agent has both inside and outside options, does the reference point established for the inside option also effect the way agents evaluate outside options? A clear answer to this question might complement the findings of this paper to shed light on which types of contracts may favor rewards or penalties.

References

Benartzi, S. and Thaler, R. "Myopic Loss Aversion and the Equity Premium Puzzle." *Quarterly Journal of Economics*, 110(1995):73 – 92.

Camerer, C. "Individual Decision Making." in John H. Kagel and Alvin E. Roth (eds.) *The Handbook of Experimental Economics.* Princeton, NJ: Princeton University Press, 1995.

Curtis, K.R. and McCluskey, J.J. "Contract Incentives in the Processed Potato Industry." Proceedings, First Biennial Conference of the Food System Research Group, June 2003.

Holmstrom, B. "Moral Hazard and Observability." *Bell Journal of Economics* 10(1979):74-91.

Horstmann, I.J.; Mathewson, F.; and Quigley, N. "Bonuses and Penalties in Incentive Contracts." Working Paper, University of Toronto Rotman School of Management, 2002.

Hueth, B. and Ligon, E. "Estimation of an Efficient Tomato Contract." *European Review of Agricultural Economics* 29(2002): 237 – 253.

Hueth, B. and Ligon, E. "On the Efficacy of Contractual Provisions for Processing Tomatoes." Working paper, Department of Agricultural and Resource Economics, University of California, Berkeley, June 11, 2003.

Innes, R.D. "Limited Liabilities and Incentive Contracting with ex ante Action Choices." *Journal of Economic Theory* 52(1990): 45-67.

Kahneman, D. and Tversky, A. "Prospect Theory: an Analysis of Decision under Risk." *Econometrica* 47(1979): 263 – 92.

Kahneman, D.; Knetsch J.; and Thaler, R. "The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives* 5(1991):193 – 206.

Lazear, E. *Personnel Economics for Managers,* New York: Wiley & Sons, Inc., 1998.

Rogerson, W. "The First-order Approach to Principal-Agent Problems," *Econometrica*, 53(1985): 1357-67.

Sandmo, A. "On the Theory of the Competitive Firm under Price Uncertainty." *American Economic Review*. 61(1971):65 – 73.

Tallitsch, J., and Moynihan, J. "Fine-Tuning Sales Compensation Programs." *Compensation Benefits Review* 26(1994):34-37.

Tversky, A. and Kahneman, D. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5(1992): 297 – 323.

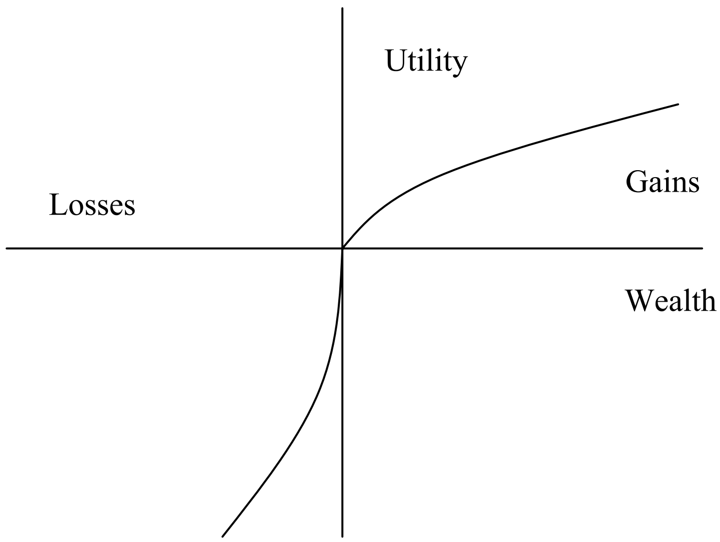

Figure 1. Prospect theoretic value function

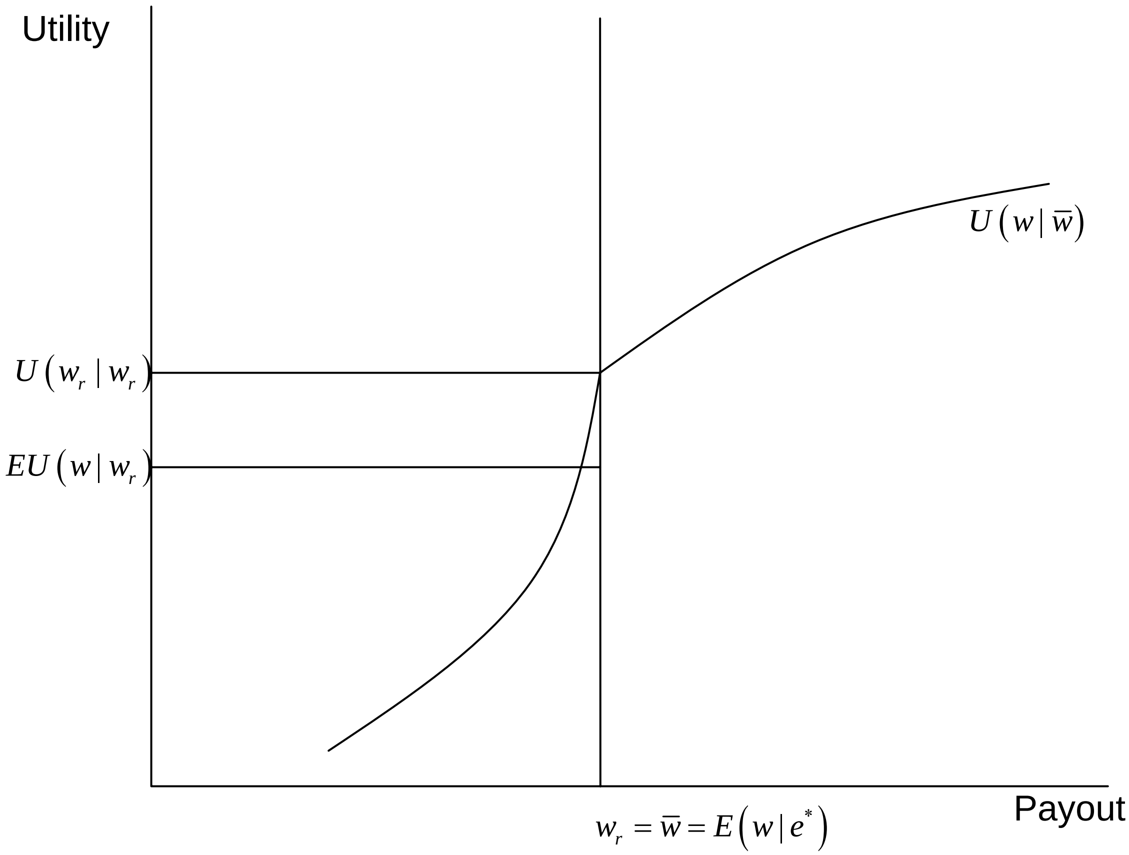

Figure 2a. Loss aversion with reservation equal to reference wage.

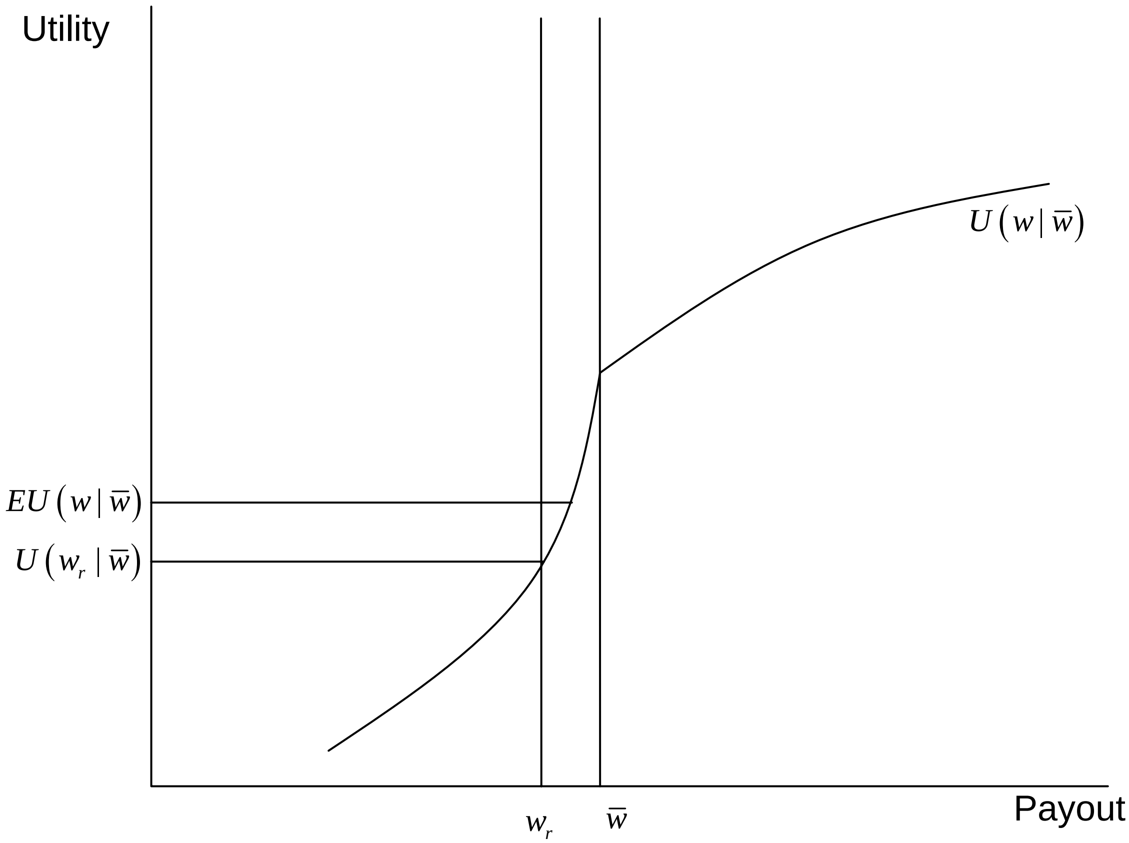

Figure 2b. . Loss aversion with reservation below reference wage.

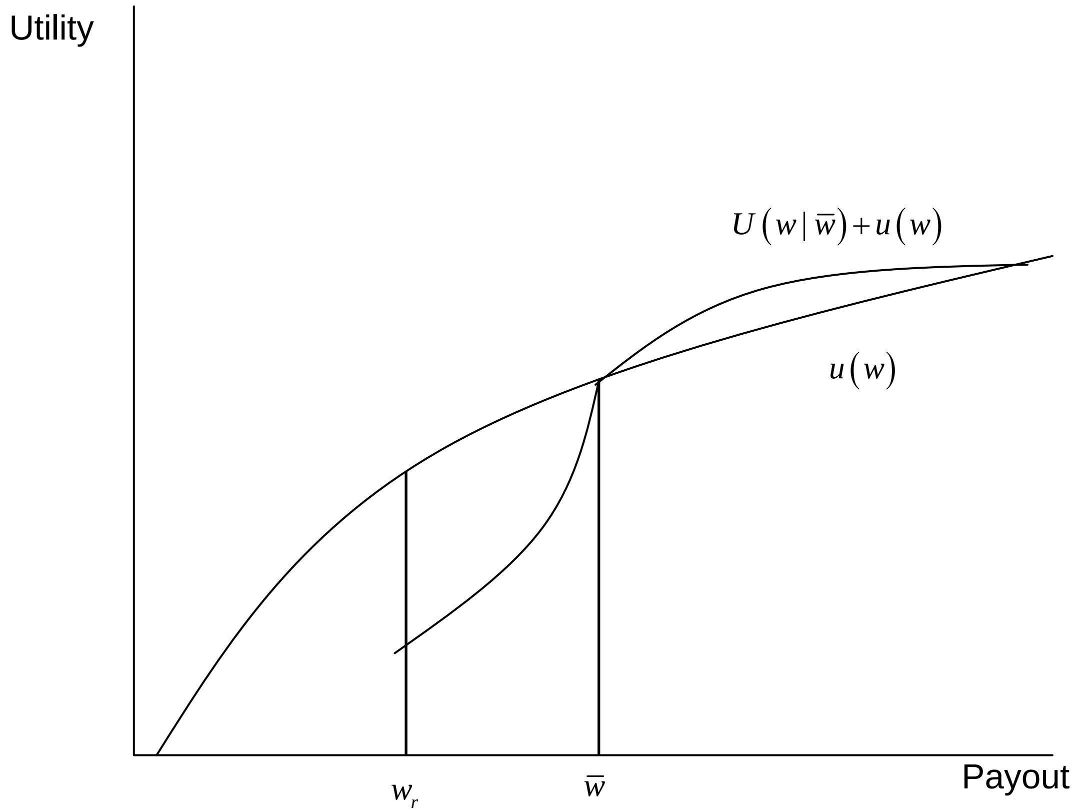

Figure 3. Loss aversion with reservation independence.

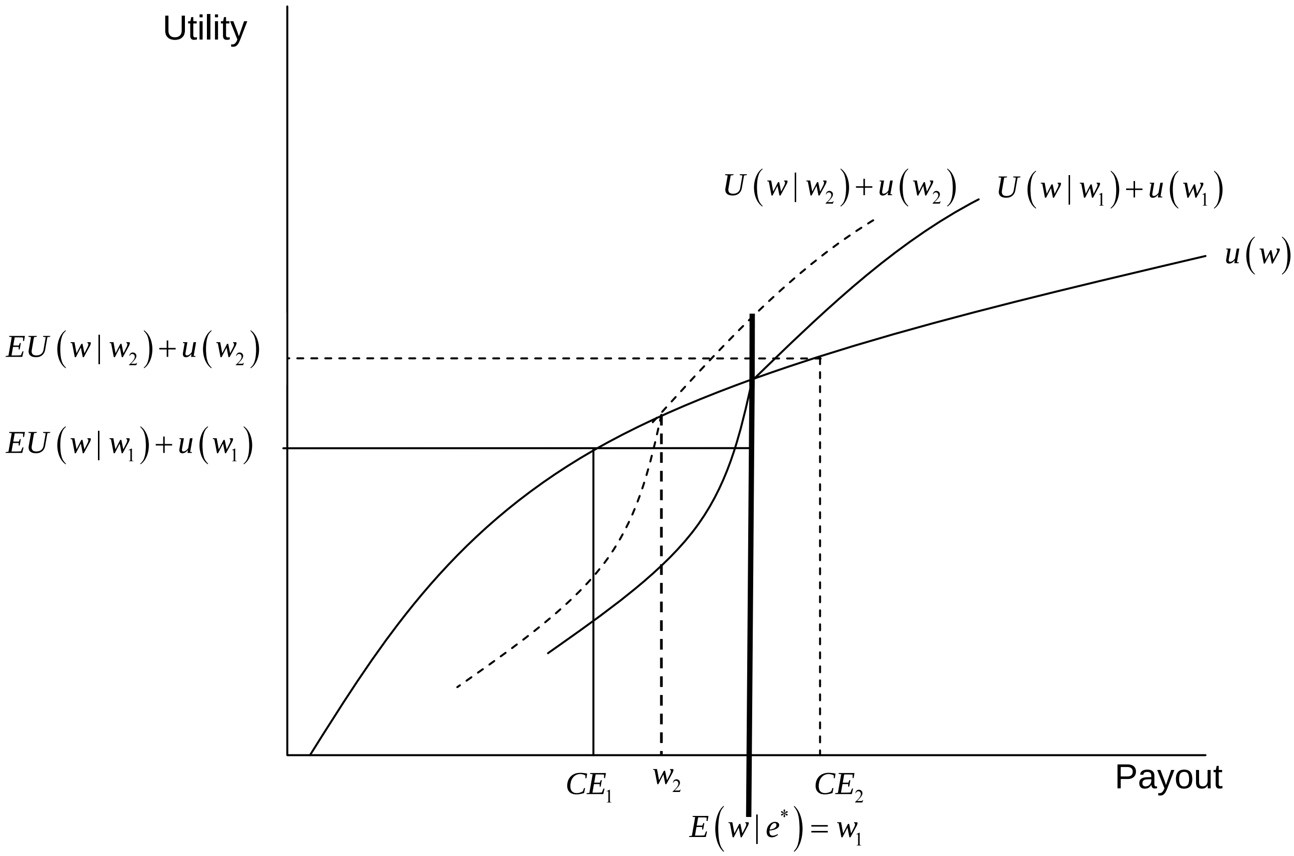

Figure 4. Certainty equivalents for reference points at the mean (w_1 , solid) and below mean (w_2 , dashed) pay under reservation independence.