

# Why Do People Pay Taxes? An Explanation Based On Loss Aversion And Overweighting of Low Probabilities

Sanjit Dhami\*      Ali al-Nowaihi†

December 2003

## Abstract

*Given actual probabilities of audit and penalty rates observed in the real world, tax evasion should be an extremely attractive gamble to an expected utility maximizer. However, in practice, one observes too much compliance relative to the predictions of expected utility. This paper considers an alternative theoretical model that is based on Kahneman and Tversky's cumulative prospect theory. The model predicts empirically plausible magnitudes of tax evasion despite low audit probabilities and penalty rates. An increase in the tax rate leads to an increase in the amount evaded- a result, which is both, intuitive, and factual, but is contrary to the prediction made by expected utility theory. Furthermore, the optimal tax rates predicted by prospect theory, in the presence of tax evasion behaviour, are consistent with actual tax rates.*

*Keywords: Tax Evasion, Loss Aversion, Decision Weights, Prospect Theory, Optimal taxation.*

*JEL Classification: D81 (Criteria for Decision Making Under Risk and Uncertainty), H21 (Efficiency; Optimal Taxation), H26 (Tax Evasion), K42 (Illegal Behavior and the Enforcement of Law).*

---

\*Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522086. Fax: +44-116-2522908. E-mail: [Sanjit.Dhami@le.ac.uk](mailto:Sanjit.Dhami@le.ac.uk).

†Department of Economics, University of Leicester, University Road, Leicester. LE1 7RH, UK. Phone: +44-116-2522898. Fax: +44-116-2522908. E-mail: [aa10@le.ac.uk](mailto:aa10@le.ac.uk).

## 1. Introduction

Why do people pay taxes? Theoretical work, largely based on expected utility theory (EUT), has struggled to cope with the stylized facts. The tax evasion puzzle goes something like this. The audit probability,  $p$ , is extremely low, with realistic magnitudes ranging from  $p = 0.01$  to  $p = 0.03$ , while the penalty rate,  $\lambda$ , that is paid in addition to repayment of the evaded tax liabilities, ranges from 0.5 to 2.0; see for example Alm et. al. (1992) and Bernasconi (1998). The seminal applications of EUT to the tax evasion problem, by Allingham and Sandmo (1972) and Yitzhaki (1974), show that the taxpayer will choose to evade taxes if the expected return on evading the tax,  $1 - p - p\lambda$ , is positive. Using fairly conservative estimates of  $p = 0.03$  and  $\lambda = 2.0$ , it turns out that the expected return on tax evasion is about 90 percent. This contradicts the empirical evidence in two ways.

1. With a positive expected return to tax evasion, EUT predicts that all taxpayers should hide some (but, typically, not all) of their income. Experimental evidence suggests that about 20 percent of tax payers fully declare all their income, even when  $p = 0$ ; see, for example, Alm et. al. (1992). Furthermore, those who decide to evade, appear to hide all their income. More accurately, tax payers, if they decide to evade, hide some sources of income completely while fully declaring other sources; see, for example Pudney et. al (2000).
2. To square the predicted extent of tax evasion under expected utility with the evidence, tax payers should be risk averse to an absurd degree; see, for example, Skinner and Slemrod (1985) and Alm et al. (1992).

But these are not the only paradoxes arising from the application of EUT to tax evasion.

3. Yitzhaki (1974) showed that under the reasonable assumption of decreasing absolute risk aversion, an increase in the tax rate leads to a decrease in tax evasion. Experimental and econometric evidence has shown this result to be counterfactual; see, for example, the references cited in al-Nowaihi and Pyle (2000).
4. Obligatory advanced tax payments should not influence the taxpayer's evasion decision under EUT. However, empirical and experimental evidence show that obligatory advanced tax payments do reduce tax evasion; see, for example, Yaniv (1999).

The Allingham-Sandmo-Yitzhaki model has been extended in a number of directions to include endogenous income, social norms, tax avoidance and uncertainty in the outcome of audits, among others; see, for example, Andreoni et. al. (1998). While these extensions

are, of course, very worthwhile, as long as EUT is retained the above paradoxes remain. Therefore, much recent work has investigated several non-EUT frameworks.

Several alternatives using non-EUT models have been proposed to explain the tax evasion puzzle. For instance, the stigma associated with being caught evading taxes might make taxpayers reluctant to evade, despite the existence of a favorable gamble; see for example Gordon (1989). Bernasconi (1998) imposes a technical restriction on preferences called “orders of risk aversion” based on earlier work by Segal and Spivak (1990)<sup>1</sup>. It turns out that for realistic  $(p, \lambda)$  pairs, if risk aversion is of order 2 then tax evasion will take place, but if it is of order 1 it might not. However, not all technical restrictions on preferences are founded on independent empirical evidence, and, hence, it is not easy to evaluate their plausibility for specific contexts and for actual behavior; see for example Starmer (2000).

On the other hand, preference restrictions can be based on a substantial body of independent experimental evidence, as for instance, in the “prospect theory” alternative to EUT given by Kahneman and Tversky (1979) and Tversky and Kahneman (1992). We apply prospect theory to the tax evasion puzzle using the latter version of prospect theory expounded in Tversky and Kahneman (1992). There are two basic foundations of prospect theory, both supported by a large body of experimental evidence.

1. Unlike EUT, where the carrier of utility is final wealth, the carriers of utility in prospect theory are gains and losses relative to some “reference point”. Experimental evidence indicates that, relative to the reference point, individuals evaluate gains and losses, differently. In particular, losses loom larger than gains: a phenomenon that in prospect theory is termed as “loss aversion”. Denote by  $x$  the payoff, relative to the reference point, in some state of the world and by  $v(x)$  the corresponding utility in that state. Loss aversion implies that the disutility arising from a loss of  $x$  pounds is larger in magnitude than the utility arising from the gain of  $x$  pounds i.e.  $v(x) < -v(-x)$ ,  $x > 0$ . Furthermore,  $v(x)$  is concave for gains, and convex for losses relative to the reference point.
2. The second foundation of prospect theory is that, unlike EUT, probabilities enter non-linearly into the evaluation of a lottery. Consistent with experimental evidence, individuals use an inverted  $S$  shaped cumulative weighting function which overweights low probabilities and underweights high probabilities.

Prospect theory has been successfully used to explain a range of puzzles in economics, such as the disposition effect, asymmetric price elasticities, elasticities of labour supply

---

<sup>1</sup>For a certain class of gambles, orders of risk aversion imposes restrictions on the first and second derivatives of the risk premium term with respect to the tax rate.

that are inconsistent with standard models of labour supply and the excess sensitivity of consumption to income; see, for example, Camerer (2000). An important application of prospect theory is to the equity-premium puzzle in finance, which is similar in spirit to the tax evasion puzzle in that both can be formulated in terms of a similar portfolio choice problem. The equity premium puzzle asks the following question. Why are annual real rates of return on stocks, relative to bonds, about 6 percent higher since 1926 when the standard deviation for the two investments equals 0.2 and 0.0 respectively? Using EUT, it is difficult to make sense of this apparently high premium on stocks, unless individuals have coefficients of relative risk aversion in excess of 30, which is unrealistic<sup>2</sup>. Benartzi and Thaler (1995) successfully use prospect theory to explain the equity premium puzzle based on loss averse agents who are also myopic in the sense of having limited planning horizons. An obvious parallel with tax evasion is to imagine evasion and compliance as two investments and ask why people invest disproportionately large amounts in the latter when returns on the former are so high?

Prospect theory has also been used to explain the tax evasion puzzle. The experimental studies of Alm et al. (1992) suggest that one possible explanation for why people pay taxes might potentially be based on prospect theory. Individuals might be using a non-linear transformation of probabilities to overweight the probability of a tax audit, which provides for an obvious deterrent to tax evasion activity, however, they do not attempt to formalize their suggestion. It is generally agreed that the most satisfactory approach of transforming probabilities is that of rank dependent expected utility; see, for example, Starmer (2000). But Eide (2001) shows that the paradoxical comparative static results of the Allingham-Sandmo-Yitzhaki model carry over to the more general rank dependent expected utility.

Yaniv (1999) applies the idea of loss aversion in prospect theory to the tax evasion problem, however, his underlying framework is very similar to expected utility theory. In particular, he uses linear weighting in probabilities while prospect theory uses a non-linear inverted *S* shaped cumulative weighting function. Furthermore, his use of “choice heuristics” to combine common terms in income arising in different states of nature is at odds with the treatment of prospect theory in Tversky and Kahneman (1992).

Our paper is closest in spirit to Bernasconi and Zanardi (2002). Bernasconi and Zanardi (2002) choose an arbitrary reference point that does not necessarily depend on the tax rate. Hence, it is possible that in both states of the world (i.e. ‘income is audited’ and ‘income is not audited’) either the tax payer is in the domain of gains or in the domain of losses or in both domains. It is, however, counterintuitive that the reference point of the taxpayer,

---

<sup>2</sup>Mankiw and Zeldes (1991) calculate that a coefficient of relative risk aversion of 30 implies that the certainty equivalent of the prospect - win 50,000 or 100,000 each with probability 1/2 is only 51,209. This is an absurdly high degree of risk aversion.

when the tax evasion decision is made, does not depend on the legal tax liabilities<sup>3</sup>. Thus, for instance, if the tax payer is always in the domain of gains, then the analysis is exactly as under expected utility and the tax evasion puzzle remains outstanding. It is only in the other two complementary cases i.e. when the taxpayer is always in the domain of losses or in both domains in different states of the world, that prospect theory has any bite in their analysis. Furthermore, they do not use cumulative transformations of probability and so it is likely that the optimal choice could be stochastically dominated by other options; see for instance, Starmer (2000).

Stigma from the detection of tax evasion seems to be recognized as an important factor in the tax evasion literature; see, for example, Slemrod and Skinner (1985). In related contexts, Besley and Coate (1992) also argue for the importance of stigma as a feature of actual behavior. Hence, we integrate stigma from tax evasion within prospect theory. Including stigma does not change the comparative static results of our model, but it allows us to predict the level of aggregate tax evasion.

We broadly ask the following questions. Is it possible to resolve the tax evasion puzzles, (1)-(4) above, using prospect theory? What role does stigma play in partitioning the population into those who pay and those who evade taxes? Furthermore, is the prospect theory framework amenable to optimal tax analysis? and, if yes, are the magnitudes of optimal tax rates predicted by prospect theory realistic<sup>4</sup>?

The results are as follows. Empirically relevant combinations of  $(p, \lambda)$  lead to realistic patterns of tax evasion when one applies cumulative prospect theory to the tax evasion decision. On the other hand, the predictions of EUT are substantially at odds with the evidence. Under prospect theory, an increase in the tax rate leads to an increase in the amount evaded, which is intuitive and factual; this is in direct contrast to the predictions made by EUT. Finally, the optimal tax rates predicted by prospect theory, in the presence of tax evasion behavior, are consistent with actual tax rates.

The plan of the paper is as follows. Section 2 describes the basic model and provides a brief overview of the method of prospect theory. Section 3 solves for the equilibrium tax evasion under EUT and prospect theory; it also provides a diagrammatic interpretation of the results. Section 4 examines the predictions of expected utility and prospect theory with respect to the audit probabilities and the penalty rates. Section 5 derives the optimal tax rate under prospect theory. Finally, Section 6 concludes. All proofs are collected in the appendix.

---

<sup>3</sup>For instance, the old and often quoted adage in public finance- “an old tax is no tax”- seems to have an obvious implication that the existing level of tax liabilities conditions ones post-tax reference income.

<sup>4</sup>To the best of our knowledge, this is the first time that this question has been considered within a prospect theory framework.

## 2. The Model

Each taxpayer has some exogenous taxable income  $W$  and can choose to declare some amount  $D \in [0, W]$ . The government levies a tax at some constant marginal rate  $t$  on declared income. An exogenous fraction  $p \in [0, 1]$  of the taxpayers are audited, subsequent to the filing of tax returns, and the audit reveals the true taxable income. If caught, the dishonest taxpayer must pay the outstanding tax liabilities  $t(W - D)$  and a penalty  $\lambda t(W - D)$  where  $\lambda > 0$  is the penalty rate. If evasion is discovered, the taxpayer also suffers some stigma  $s(W - D)$  where  $s$  is the stigma rate on evaded income<sup>5</sup>. There is a continuum of individuals with stigma rates  $s \in \left[ \underline{s}, \bar{s} \right]$  with density  $\phi(s)$  and a distribution function  $\Phi(s)$ . Denoting by  $Y_A$  and  $Y_{NA}$  respectively, the income of the taxpayer when he is audited and when he is not,

$$Y_{NA} = W - tD, \text{ and} \quad (2.1)$$

$$Y_A = W - tD - t(1 + \lambda)(W - D) - s(W - D). \quad (2.2)$$

### 2.1. Sequence of Moves

The tax authorities move first, making an announcement of  $p$  and  $\lambda$ . The taxpayer then makes the decision to either fully report his or her income or evade a fraction of it. Finally, the government audits a fraction  $p$  of the returns and dishonest taxpayers are required to give up a fraction  $(1 + \lambda)$  of their unreported income.

### 2.2. The Decision Problem Under Expected Utility Theory

Denote the set of outcomes, possibly income levels, in a gamble as  $Y_1, \dots, Y_n$ . Under EUT, the carrier of utility is the final wealth level and not, unlike prospect theory, wealth relative to some reference income level. In order to facilitate comparison with the results under prospect theory, the utility of any outcome is defined as follows.

$$v(Y_i) = \begin{cases} Y_i^\beta & \text{if } Y_i \geq 0 \\ 0 & \text{if } Y_i < 0 \end{cases} \quad (2.3)$$

---

<sup>5</sup>As in Gordon (1989) and Besley and Coate (1992), such stigma enters linearly, as a monetary equivalent, into the payoff in that state of the world. Although, a natural interpretation of stigma might include factors such as the tug on one's conscience, and loss of face among family and community etc., other interpretations of stigma are possible. These might include, in an appropriately specified dynamic game, the effect on ones current and future earnings arising from the bad publicity one receives from tax evasion.

In (2.3) the parameter  $\beta \in (0, 1)$ <sup>6</sup>. Given  $p$  and  $\lambda$  announced by the tax authorities, the tax payer chooses declared income  $D \in [0, W]$  in order to maximize expected utility, which, using 2.1 and 2.2, is given as

$$E[U] = \begin{cases} (1-p)[Y_{NA}]^\beta + p[Y_A]^\beta & \text{if } Y_A \geq 0 \\ (1-p)[Y_{NA}]^\beta & \text{if } Y_A < 0 \end{cases} \quad (2.4)$$

### 2.3. Utility of an Outcome Under Prospect Theory

Denote the reference income of the taxpayer as  $R$  and the income relative to the reference point as  $X_i = Y_i - R$ ,  $i = 1, \dots, n$ . As in Kahneman and Tversky (1979) and Tversky and Kahneman (1992), the utility associated with any outcome  $X_i$  is defined as  $v(X_i)$ , and

$$v(X_i) = \begin{cases} X_i^\beta & \text{if } X_i \geq 0 \\ -\theta(-X_i)^\beta & \text{if } X_i < 0 \end{cases} \quad (2.5)$$

where  $\beta \in [0, 1]$  and  $\theta > 1$  are preference parameters. Based on experimental evidence Kahneman and Tversky (1979) and Tversky and Kahneman (1992) suggests that  $\beta \cong 0.88$  while  $\theta \cong 2.25$ .

### 2.4. The Reference Point Under Prospect Theory

Although prospect theory does not provide sufficient guidance to determine the reference point in each possible situation, in several cases there is an obvious candidate for a reference point. We take the legal after-tax income,  $(1-t)W$ , as the reference point in this paper using the old adage in public finance that “an old tax is no tax”<sup>7</sup>.

### 2.5. The Decision Problem Under Prospect Theory

A brief sketch of the generic formulation under prospect theory is given below; details can be seen in Tversky and Kahneman (1992). There are two possible states of the world in the model, defined by  $\Lambda$  such that

$$\Lambda = \{Audit, No\ Audit\}.$$

The “audit state” occurs with probability  $p$  and the “no audit state” occurs with probability  $1-p$ . Relative to the reference income, in the “audit state” the taxpayer’s income is  $X_A = Y_A - (1-t)W = -[\lambda t + s](W - D) \leq 0$ , while in the “no audit state” it is  $X_{NA} = Y_{NA} - (1-t)W = t(W - D) \geq 0$ , therefore,

---

<sup>6</sup>From (2.2) it is possible that  $Y_A < 0$ . The simplest way to incorporate  $Y_A < 0$  into EUT is to define  $v(Y_A) = 0$ , as in (2.3). Other possibilities are  $v(Y_A) = -v(-Y_A)$  and  $v(Y_A) = -\infty$ . None of our results depend on which specification we use for  $v(Y_A)$  when  $Y_A < 0$ .

<sup>7</sup>The implications of a reference income other than  $(1-t)W$  are taken up in the appendix.

$$X_{NA} = t(W - D), \text{ and} \quad (2.6)$$

$$X_A = -[\lambda t + s](W - D). \quad (2.7)$$

Define the set  $X$  as the set of monetary consequences, hence,

$$X = \{-[\lambda t + s](W - D), 0, t(W - D)\}$$

The “neutral outcome” 0 is also included in  $X$  and all other outcomes are measured as gains or losses relative to it. Consider the following prospect, denoted by  $f$ .

$$f = \{-[\lambda t + s](W - D), p; t(W - D), 1 - p\}$$

In this literature,  $f^+$  and  $f^-$  respectively denote the positive and negative parts of the prospect<sup>8</sup>. To each possible prospect  $f$ , prospect theory assigns a number  $V(f)$  such that prospect  $f$  is preferred to prospect  $g$  if and only if  $V(f) > V(g)$ . The “value equation” in prospect theory, which the individual tries to maximize, is very similar in form to the rank expected utility models and is given by

$$V(f) = \pi_i^+ v(t(W - D)) + \pi_i^- v(-[\lambda t + s](W - D)), \quad (2.8)$$

The payoff arising from any individual outcome,  $v(X_i)$ , is defined in (2.5). Unlike EUT, the outcomes are not simply weighted by objective probabilities but rather they are weighted by some “decision weights”  $\pi_i$  which are non-linear transformations of objective probabilities. Furthermore, the positive and negative parts of the prospect are generally separately weighted by different weighting functions denoted by  $\pi^+$  and  $\pi^-$  respectively.

The decision weights are transformations of cumulative probabilities<sup>9</sup>, as in rank dependent models of expected utility. In particular, at the highest (respectively lowest)

---

<sup>8</sup>In the tax evasion case,  $f^+ = \{0, p; t(W - D), 1 - p\}$  and  $f^- = \{[\lambda t + s](W - D), p; 0, 1 - p\}$ , hence,  $f = f^+ - f^-$ . Notice that the neutral outcome is included in each of the prospects because it is neither positive nor negative. The distinction between  $f^+$  and  $f^-$  is made because the probability weighting functions for the negative and positive parts of the prospect are typically distinct; see Tversky and Kahneman (1992) for the details.

<sup>9</sup>In the more general case suppose that there are  $n$  possible outcomes  $X_1, \dots, X_n$  in the domain of gains with associated objective probabilities  $p_1, \dots, p_n$ . Then the decision weight  $\pi_i^+(X_i) = w^+(p_i + \dots + p_n) - w^+(p_{i+1} + \dots + p_n)$  while  $\pi_n^+(X_n) = w^+(p_n)$ . The function  $w : [0, 1] \rightarrow [0, 1]$  is a weighting function that transforms objective probabilities into subjective weights. In the domain of cumulative probabilities, Tversky and Kahneman (1992) show that the shape of the weighting function consistent with experimental evidence is an inverted  $S$  shape.

Denoting the  $m$  possible outcomes in the domain of losses  $X_{-m}, \dots, X_{-1}$  with associated probabilities  $p_{-m}, \dots, p_{-1}$  one defines  $\pi_i^-(X_i) = w^-(p_{-m} + \dots + p_i) - w^-(p_{-m} + \dots + p_{i-1})$  with  $\pi_{-m}^-(X_{-m}) = w^-(p_{-m})$ . The decision weights in the paper now follow from noting that  $n = m = 1$  for the tax evasion case.



value of the outcome in the prospect  $f^+$  (respectively  $f^-$ ),  $\pi_i^+ = w^+(1-p)$  (respectively  $\pi_i^- = w^-(p)$ ), where  $w^+$  and  $w^-$  are “weighting functions” that apply respectively to the positive and the negative parts of the prospect; these are formally defined below. Hence, in the special case of there being only one outcome each in the domain of gains and losses, as is the case in the generic tax evasion problem, the decision weights based on cumulative probabilities are simply equal to weights based on objective probabilities. Using the definition of  $v(X_i)$  in 2.5, this allows the expression 2.8 to be written as

$$V(f) = w^+(1-p) [t(W-D)]^\beta + w^-(p)(-\theta) [(\lambda t + s)(W-D)]^\beta \quad (2.9)$$

Comparing the value equation in (2.8) with the expression for expected utility in (2.4), aside from the difference that the carriers of value in prospect theory are gains and losses relative to some reference point rather than final wealth levels, the two main differences are the following. First, one uses decision weights in prospect theory to aggregate outcomes while one uses direct objective probabilities under EUT. Second, there is loss aversion (arising through the term  $-\theta$ ) in prospect theory while there is none under EUT.

## 2.6. The Weighting Function in Prospect Theory

Empirical evidence is widely consistent with an inverted  $S$  shaped form for the weighting function; see for example Kahneman and Tversky (1979), Tversky and Kahneman (1992) and Prelec (1998). Denoting by  $\mathbf{p}$  the cumulative probability, Prelec (1998) derives the following weighting function,

$$w(\mathbf{p}) = w^+(\mathbf{p}) = w^-(\mathbf{p}) = \exp [ -(-\ln \mathbf{p})^\alpha ], \quad (2.10)$$

where  $0 < \alpha \leq 1$ . There are several advantages of this weighting function, relative to the others suggested in the literature<sup>10</sup>. First, it has an inverted  $S$  shape which is consistent with experimental evidence. Second, it is based on axiomatic foundations. Third, it has the same form for gains and losses. For these reasons, we choose to use this weighting function in this paper. Several other weighting functions have been considered in the literature, that are based on transformations of individual (and not cumulative) probabilities, however, as is well known, in decision weighted models, these lead to the

---

<sup>10</sup>Tversky and Kahneman (1992) suggest the following weighting function, based on experimental evidence, in the domain of gains and losses respectively

$$w^+(\mathbf{p}) = \mathbf{p}^\gamma \{ \mathbf{p}^\gamma + (1-\mathbf{p})^\gamma \}^{-1/\gamma}; \quad \gamma \cong 0.61 \quad (2.11)$$

$$w^-(\mathbf{p}) = \mathbf{p}^\delta \{ \mathbf{p}^\delta + (1-\mathbf{p})^\delta \}^{-1/\delta}; \quad \delta \cong 0.69 \quad (2.12)$$

violation of first order stochastic dominance and for this reason these are not attractive<sup>11</sup>; see, for example, Starmer (2000).

### 3. The Tax Evasion Decision

This section derives the optimal tax evasion decision under expected utility and prospect theory, analytically, as well as diagrammatically.

#### 3.1. The Tax Evasion Decision Under Expected Utility

Differentiating (2.4) in the region  $Y_A > 0$ , the first order condition to the tax evasion problem is given by

$$\frac{\partial E[U]}{\partial D} = -t\beta(1-p)[Y_{NA}]^{\beta-1} + p\beta[t\lambda + s][Y_A]^{\beta-1} \leq 0; \quad D \geq 0 \quad (3.1)$$

where  $Y_A$  and  $Y_{NA}$  are defined in 2.2 and 2.1 respectively. The second order condition is given by

$$\frac{\partial^2 E[U]}{\partial D^2} = -t^2\beta(1-\beta)(1-p)[Y_{NA}]^{\beta-2} - \beta(1-\beta)p[t\lambda + s]^2[Y_A]^{\beta-2} < 0. \quad (3.2)$$

From 3.2 it follows that the necessary and sufficient condition for the maximization of expected utility in the region  $Y_A > 0$ , is

$$\begin{aligned} \frac{\partial E[U]}{\partial D} &\geq 0 \quad \text{for } D = W \\ \frac{\partial E[U]}{\partial D} &= 0 \quad \text{for } 0 < D < W \\ \frac{\partial E[U]}{\partial D} &\leq 0 \quad \text{for } D = 0 \end{aligned} \quad (3.3)$$

If the taxpayer does not evade ( $D = W$ ) then  $Y_A = W - tD > 0$  and, hence, (3.1), (3.2), and the first row of (3.3) all apply. In particular,  $\frac{\partial E[U]}{\partial D} \geq 0$  gives  $-t(1-p) + p[t\lambda + s] \geq 0$  and, hence,

$$\lambda(p) \geq \frac{1-p}{p} - \frac{s}{t} \quad (3.4)$$

In  $(\lambda, p)$  space, (3.4) gives the set of  $(\lambda, p)$  points such that a taxpayer with stigma  $s \in \left[ \underline{s}, \bar{s} \right]$  reports the full amount of income. At all point below this set, the taxpayer

---

<sup>11</sup>An example of this type of function is the one proposed by Camerer and Ho (1994)  $f(p) = 1 - (1-p)^\gamma / \{p^\gamma + (1-p)^\gamma\}^{1/\gamma}$  with  $f(0.7) \approx 0.7$ . For an application of this sort of weighting function to an essentially expected utility framework of tax evasion, see Bernasconi (1998).

chooses to evade some strictly positive fraction of income. Using (3.4), in the benchmark case of  $s = 0$ , the taxpayer will evade some fraction of income when  $1 - p - p\lambda > 0$ , as claimed in the introduction.

Solving (3.1) for an interior solution to the declared income, one obtains

$$D = W \left[ 1 - (1 - t) \left[ (t\lambda + s) + t \left( \frac{(t\lambda + s)p}{t(1 - p)} \right)^{\frac{1}{1-\beta}} \right]^{-1} \right]. \quad (3.5)$$

Implicit differentiation of the first order condition (3.1), or direct differentiation of (3.5) gives rise to all the well known results under EUT, in particular, that the declared income is increasing in  $p$  and  $\lambda$ . Since stigma  $s$  plays a role that is identical to  $\lambda$ , it is easy to check that declared income is also increasing in  $s$ . It is also straightforward to derive the result in Yitzhaki (1974), namely that the declared income increases with the tax rate if preferences exhibit decreasing absolute risk aversion<sup>12</sup>. These results are summarized, without proof, in Proposition 1 below.

**Proposition 1** : *Under EUT, tax evasion is decreasing in  $p$ ,  $\lambda$ , and  $s$ . If preferences exhibit decreasing absolute risk aversion, then tax evasion is decreasing in the tax rate.*

### 3.2. The Tax Evasion Decision Under Prospect Theory

The taxpayer maximizes the value equation in (2.9) by an appropriate choice of  $D$ , given the weighting function in 2.10, which weights gains and losses symmetrically; rewriting (2.9)

$$V(f) = (W - D)^\beta h(\lambda, p, s, t, \theta, \beta), \quad (3.6)$$

where

$$h = h(\lambda, p, s, t, \theta, \beta) = t^\beta \left( \exp[-(-\ln(1 - p))^\alpha] - \theta \exp[-(-\ln p)^\alpha] \left( \lambda + \frac{s}{t} \right)^\beta \right) \geq 0 \quad (3.7)$$

Since the optimization problem stated in (3.6) is linear in  $D$ , the solution is given by

$$\begin{cases} D = 0 & \text{if } h > 0 \\ D \in [0, W] & \text{if } h = 0 \\ D = W & \text{if } h < 0 \end{cases} \quad (3.8)$$

The solution to the tax evasion problem under prospect theory is a bang-bang solution<sup>13</sup>. The solution is shown in a diagram below that also clarifies the intuition behind the result.

<sup>12</sup>These results are widely available and so no derivations are presented; see for example Myles (1995).

<sup>13</sup>By using other reference incomes, which are not as obvious and plausible as the legal after-tax income,

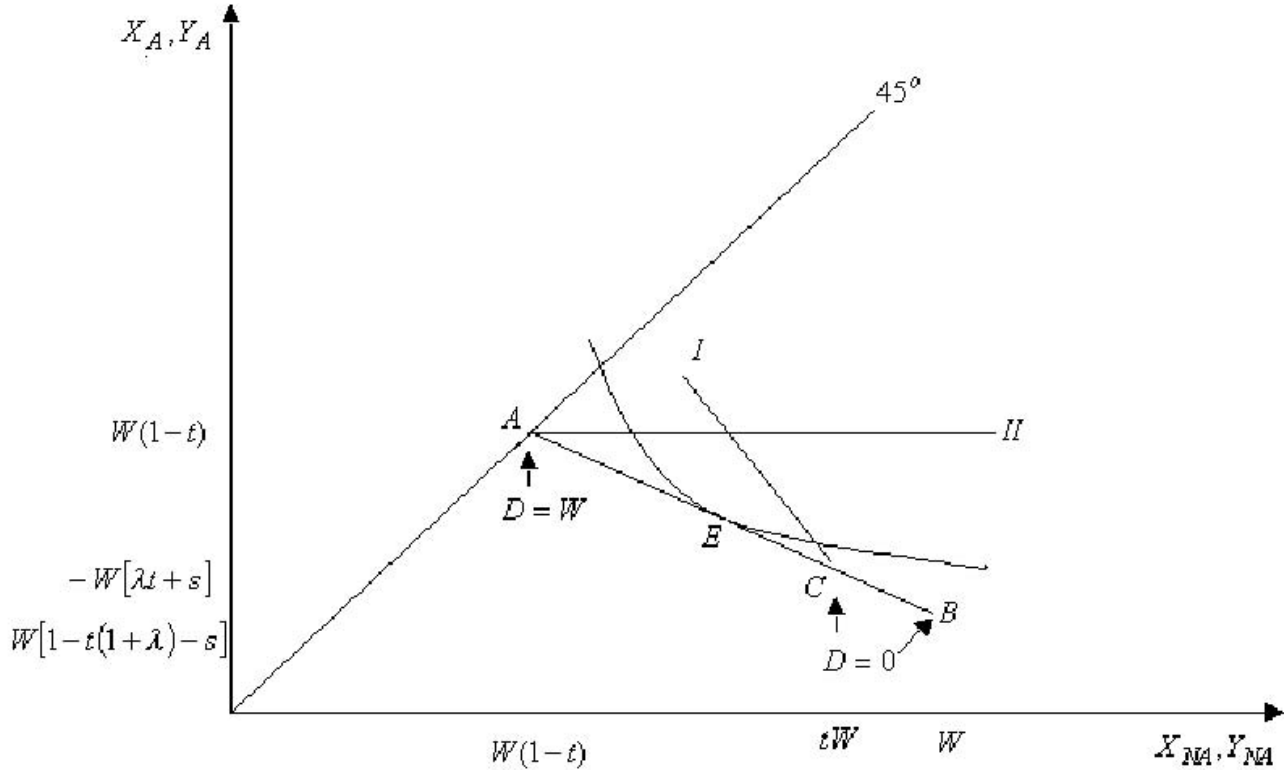


Figure 3.1: Tax evasion under expected utility and prospect theory

### 3.3. A Diagrammatic Representation of the Solution in State Space

It is instructive to compare the solutions under EUT and prospect theory in the space of incomes in the two states of the world, “audit” and “no audit”; this is done in Figure 3.1 below. Under EUT, the carriers of wealth are final wealth levels, so the origin corresponds to  $(0, 0)$ . Under prospect theory, however, the carriers of utility are gains and losses relative to the reference point,  $W(1-t)$ , hence, the origin is at  $(W(1-t), W(1-t))$ , which corresponds to point  $A$  in the figure. Therefore, the diagram for expected utility is drawn in  $(Y_A, Y_{NA})$  space with origin at  $(0, 0)$  while the diagram for prospect theory is drawn in  $(X_A, X_{NA})$  space with origin at  $[W(1-t), W(1-t)]$ .

#### 3.3.1. Tax Evasion Under Expected Utility

$(1-t)W$ , one can generate interior solutions; this is shown in the appendix. An advantage of the bang-bang solution, however, is that it is very amenable to analysis and also descriptive of several forms of tax evasion which take the form of hiding certain activities completely from the tax authorities; see, for example, Pudney et. al. (2000).

In Figure 3.1, point  $A$  corresponds to the case where the taxpayer declares all income so  $D = W$ . Using (2.1) and (2.2), it can be checked that the coordinates of this point are  $W(1 - t), W(1 - t)$ . Point  $B$  at the other extreme, corresponds to the case where  $D = 0$ ; again using (2.1) and (2.2), it can be checked that the coordinates of this point are  $W, W [1 - t(1 + \lambda) - s]$ . Hence, the slope of the budget constraint  $AB$ , any point of which reflects a tax evasion decision  $D \in [0, W]$ , can be checked to equal  $-\left(\lambda + \frac{s}{t}\right)$ . Using (2.4), the slope of an indifference curve under EUT is given by  $\frac{dY_A}{dY_{NA}} = \frac{-(1-p)}{p} \frac{v'(Y_{NA})}{v'(Y_A)}$ . Along a ray of  $45^\circ$  from the origin where  $Y_A = Y_{NA}$  the slope of the indifference curve is  $\frac{dY_A}{dY_{NA}} = \frac{-(1-p)}{p}$ .

Whether the individual evades any tax or not depends on whether, at point  $A$ , the indifference curve is steeper or flatter than the budget constraint. The diagram shows the case where the former is true so that at the equilibrium point for the taxpayer under EUT,  $E$ , some strictly positive tax evasion takes place.

The comparative static results with respect to  $s$  and  $\lambda$  are now straightforward. An increase in  $s$  or  $\lambda$  lowers point  $B$  and so the budget constraint swivels down while being anchored at point  $A$  but the slope of the indifference curve along the  $45^\circ$  line is unchanged. Hence, by making the budget constraint steeper, an increase in  $s$  or in  $\lambda$  reduces evasion by pushing the equilibrium point  $E$  in the north-west direction. An increase in the tax rate, on the other hand, shifts the entire budget constraint inwards and also alters its slope. It can be checked that the corresponding income and substitution effects work in different directions; Yitzhaki (1974) showed that in the case of decreasing absolute risk aversion, the two effects work in the same direction, namely an increase in declared income<sup>14</sup>.

### 3.3.2. Tax Evasion Under Prospect Theory

Under prospect theory, by using (2.6) and (2.7), it is easy to check that the decision  $D = W$  corresponds to  $[X_A, X_{NA}] = [W(1 - t), W(1 - t)]$  which is the origin in  $X_A, X_{NA}$  space and is denoted by the point  $A$ . Again, using (2.6) and (2.7) check that the tax evasion decision  $D = 0$ , corresponds to the coordinates  $tW, -(\lambda t + s)W$  which in the diagram is shown as point  $C$ . The slope of the budget constraint,  $AC$ , under prospect theory, then equals  $-\left(\lambda + \frac{s}{t}\right)$ , which is identical to the slope of the budget constraint,  $AB$ , under expected utility. Hence, the budget constraint  $AC$  lies on top of  $AB$  but is shorter.

Using (2.9), the slope of the indifference curve is given by  $\frac{dX_A}{dX_{NA}} = \frac{-w(1-p)}{w(p)} \frac{v'(X_{NA})}{v'(X_A)}$ . Substituting the values for  $X_A$  and  $X_{NA}$  one can check that  $\frac{dX_A}{dX_{NA}} = \frac{-w(1-p)}{w(p)} \left(\frac{1}{\theta}\right) \left(\lambda + \frac{s}{t}\right)^{1-\beta}$ . Thus, the slope of the indifference curve is independent of  $X_{NA}$  and so the indifference curves are linear which leads to corner, bang-bang, solutions. If the indifference curves are

---

<sup>14</sup>For a diagrammatic but more formal treatment see, for example, al-Nowaihi and Pyle (2000).

relatively flat (see the indifference curve marked *II*) then the taxpayer does not evade any taxes while in the complementary case (see the indifference curve marked *I*) the taxpayer does not declare any income.

To fix ideas, it is useful to think of individuals undertaking several activities, some of which cannot be evaded, while there are others in which evasion takes the form of the individual completely hiding the activity for tax purposes. Hence, tax evasion takes the form of either complete evasion of the evadable activities or complete reporting of the non-evadable activities<sup>15</sup>.

Comparing the slopes of the budget constraint and the indifference curves under prospect theory, the following is immediate

$$\left\{ \begin{array}{lll} \text{Case : I} & D = 0 & \text{if } \left(\frac{1}{\theta}\right) \frac{w(1-p)}{w(p)} > \left(\lambda + \frac{s}{t}\right)^\beta \\ \text{Case : II} & D \in [0, W] & \text{if } \left(\frac{1}{\theta}\right) \frac{w(1-p)}{w(p)} = \left(\lambda + \frac{s}{t}\right)^\beta \\ \text{Case : III} & D = W & \text{if } \left(\frac{1}{\theta}\right) \frac{w(1-p)}{w(p)} < \left(\lambda + \frac{s}{t}\right)^\beta \end{array} \right. \quad (3.9)$$

The diagrammatic equilibrium derived in (3.9) is identical to the analytical equilibrium in (3.8). It is clear from the inequality given in Case III of (3.9), that in prospect theory, the following factors are conducive to eliminating tax evasion, i.e. for  $D = W$ .

1. Low tax rates,  $t$ .
2. High levels of stigma,  $s$ .
3. High penalties for tax evasion,  $\lambda$ .
4. High levels of loss aversion,  $\theta$ .
5. Overweighting of the probability of a loss,  $p$ .

The comparative static properties of the tax evasion model under prospect theory are formally summarized below.

**Proposition 2** : Given  $\lambda, s, p, \theta$ , and  $\beta$ ,  $\exists t = t_c \in [0, 1]$  such that the individual does not evade taxes if  $t < t_c$  but evades taxes if  $t > t_c$ . Given  $t, s, p, \theta$ , and  $\beta$ ,  $\exists \lambda = \lambda_c \geq 0$ , such that, when  $\lambda > \lambda_c$ , the taxpayer declares the full amount of income, while in the complementary case, income is evaded. Given  $t, \lambda, p, \theta$ , and  $\beta$ ,  $\exists s = s_c \in \left[ \underline{s}, \bar{s} \right]$ , such that, when  $s > s_c$ , the taxpayer declares the full amount of income, while in the complementary case, income is evaded.

<sup>15</sup>The empirical evidence presented in Pudney et. al. (2000) is suggestive of this form of tax evasion. For instance, they find that there is no evidence of a disincentive impact of extra punishment on those who decide to evade, which would, for instance, be the case if one is at a corner solution.

Comparing Propositions 1 and 2 it turns out that EUT and prospect theory both predict that tax evasion decreases with an increase in  $p$ ,  $\lambda$ , and  $s$ . However, the main difference in the two approaches is in the predicted effect of the tax rate on the amount evaded. EUT predicts, as in Yitzhaki (1974), that (with decreasing absolute risk aversion) individuals evade less income as the tax rate increases while prospect theory predicts the more intuitive and factual result that tax evasion increases with an increase in the tax rate.

#### 4. Audit Probabilities, Penalty Rates And Tax Evasion

What accounts for the high degree of tax compliance when actual  $(\lambda, p)$  combinations are such that  $1 - p - p\lambda > 0$ ? To answer this question, it is useful to start with the following question. Given reasonable magnitudes of tax evasion in the population, what combinations of  $(\lambda, p)$  are required to sustain that level of evasion. Since the weighting function in prospect theory is highly non-linear, analytical answers are not possible. Hence, the analysis below provides some simulation results in a calibrated model.

From estimates in Tversky and Kahneman (1992) we know that (approximately)  $\beta = 0.88$  and  $\theta = 2.25$ . We use a tax rate of 30 percent i.e.  $t = 0.3$ . From Bernasconi (1998) we know that in actual practice 30 to 60 percent of taxpayers report their incomes correctly. Since these figures would include those who unintentionally report incorrectly we use the more conservative estimate that only about 30 percent evade taxes. We assume that the stigma rate is distributed uniformly over the unit interval i.e.  $s \in [0, 1]$  so that, in monetary terms, stigma costs do not exceed the value of the income evaded. Given these assumptions we use the cutoff value for stigma  $s = 0.3$  which under the assumptions corresponds to an evasion rate of about 30 percent of the population; given the form of the  $(\lambda, p)$  locus in our model, the assumptions on the stigma rates do not have any great deal of significance. The question that we ask then is the following. Under this set of conditions, are the  $(\lambda, p)$  combinations predicted by EUT and prospect theory, in accord with the observed penalty and audit rates?

Under expected utility then, using 3.4, the locus of  $(\lambda, p)$  combinations that need to be consistent with the actual data are given by

$$\lambda^{EU}(p) = \frac{1-p}{p} - \frac{0.3}{0.3} \quad (4.1)$$

Using 3.9, a similar locus of  $(\lambda, p)$  combinations under prospect theory is given by

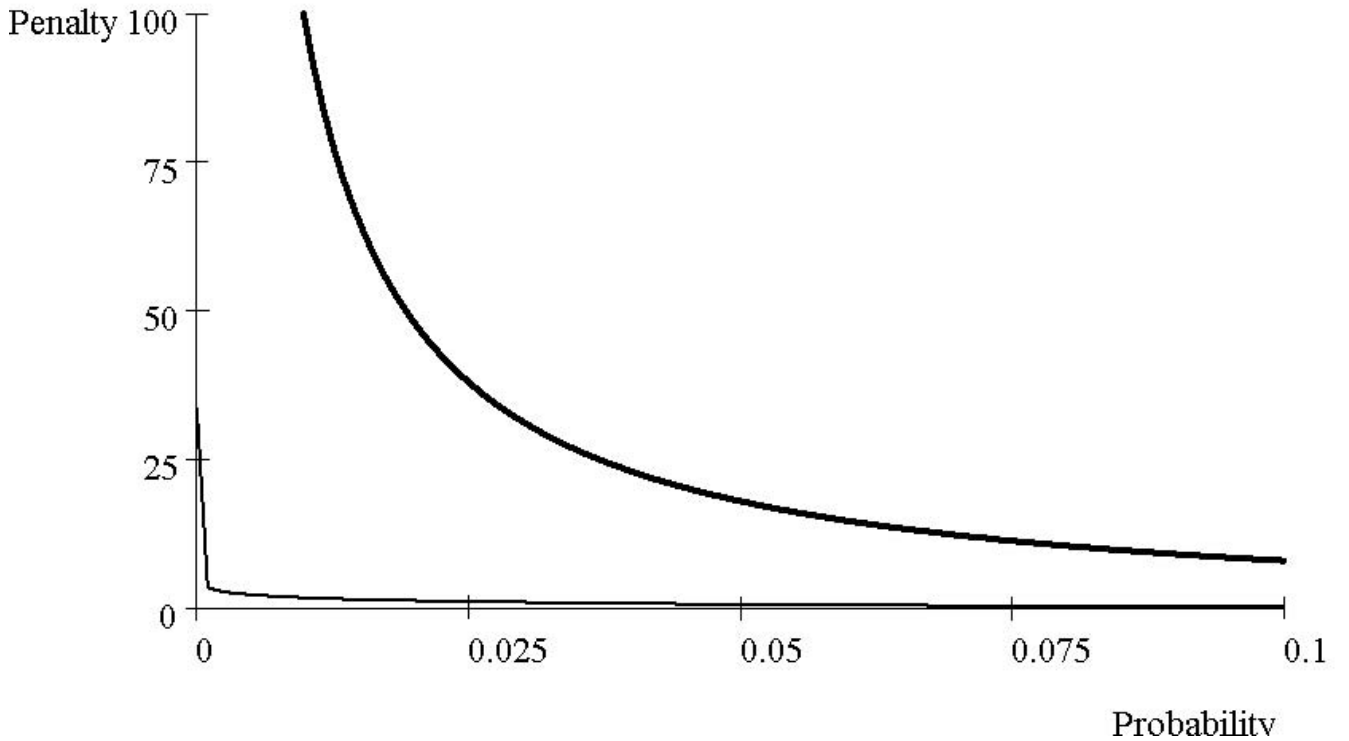
$$\lambda^{PT}(p) = \left( \frac{1}{2.25} \right)^{\frac{1}{0.88}} \Gamma(p)^{\frac{1}{0.88}} - \frac{0.3}{0.3}, \quad (4.2)$$

where

$$\Gamma(p) = \exp [(-\ln p)^{0.35} - (-\ln(1 - p))^{0.35}]$$

is obtained by using the weighting function in (2.10) and by substituting the value of  $\alpha = 0.35$ .

In the diagram below, the horizontal axis represents the audit probability  $p$  and the vertical axis represents the audit penalty,  $\lambda$ . Since most realistic audit probabilities are in the range 0.01 to 0.03 the horizontal axis is shown up to a maximum value of 0.10. The upper, thicker, curve plots the  $(\lambda, p)$  locus for expected utility in (4.1) and the lower, thinner, curve plots the  $(\lambda, p)$  locus for prospect theory in (4.2). It is clear that the  $(\lambda, p)$  locus for EUT is everywhere significantly above that for prospect theory.



Comparison of  $\lambda, p$  combinations consistent with magnitudes of actual tax evasion.

Even more telling are the numerical magnitudes of  $(\lambda, p)$  that correspond to the two loci; these are shown in the table below.

**Table I: Comparison of  $\lambda$  Under Expected Utility and Prospect Theory**

$p$	.005	.010	.015	.020	.025	.030	.035	.040	.045	.050
$\lambda^{EU}(p)$	197.67	97.67	64.33	47.67	37.67	31.00	26.23	22.67	19.89	17.67
$\lambda^{PT}(p)$	1.55	1.21	1.00	0.86	0.75	0.66	0.58	0.52	0.46	0.41

For most realistic audit probabilities i.e.  $p \in [0.01, 0.03]$  the estimate of the penalty rate ranges from 0.66 to 1.21 for Prospect Theory. This is consistent with actual values for the



audit probabilities which range from 0.5 in USA to 2.0 in Italy; see Bernasconi (1998) for these figures. On the other hand EUT predicts a penalty rate that is on average about 100 times larger in this range. This provides strong vindication of the choice of prospect theory in explaining actual parameters of policy choice relevant for the tax evasion problem<sup>16</sup>. It is remarkable that the predicted magnitudes of  $(\lambda, p)$  are so close to the actually observed values when one considers that the parameter values used for the calibration exercise were obtained in independent experimental evidence in generic situations of risk.

## 5. Optimal Choice of the Tax Rate

Tax paying authorities typically have limited choice over both  $p$  and  $\lambda$ . The constitutions in most Western democracies require that comparable offenses must be punished in a comparable manner, hence, the penalty rates levied by the tax authorities are severely constrained by penalties for similar cases of fraud and negligence. Although, in principle, the tax authorities have more flexibility in determining the audit probability, significant alternations in  $p$  are fairly resource intensive and typically the tax authorities face fairly stringent resource constraints in this endeavour. Hence, the audit probabilities in actual practice tend to lie between 0.01 and 0.03 for most countries. However, there is a much greater element in the choice of a tax rate, an issue which this section turns to, below.

In actual practice, the objective function of a government over the tax rate can be fairly complex and it can involve trade-offs between efficiency and equity as well as political economy considerations. However, the exposition below expositis the optimal choice of tax rate under prospect theory, when, following an old lineage in tax policy, the government acts like a revenue maximizing leviathan, whose sole objective is to maximize total tax revenues. It turns out that even in this fairly conservative framework, for plausible parameter values, one gets realistic magnitudes of optimal tax rates using prospect theory.

Using (3.9) one can find cutoff values of stigma  $s^*$ , such that all individuals with  $s^* < s$  choose to declare their (full) income, while all individuals with  $s < s^*$  choose to evade the (full amount of) tax liability. It can be checked, using (3.9), that

$$s^* = t\Psi; \text{ where } \Psi = \left(\frac{1}{\theta}\right)^{\frac{1}{\beta}} [\exp((-\ln p)^\alpha - (-\ln(1-p))^\alpha)]^{\frac{1}{\beta}} - \lambda. \quad (5.1)$$

The government's objective is to choose its optimal tax rate  $t^*$  in order to maximize the tax revenue  $T$ .

---

<sup>16</sup>It is worth pointing out that the results are fairly robust to actual values of  $s$  used in the calibration exercise. In each of the locii given in (4.1) and (4.2), the last term is  $\frac{s}{t} = \frac{0.3}{0.3} = 1$ , which reduces the penalty rate by only one unit. Since the actual percentage of tax evaders in Western democracies is unlikely to exceed 30 percent so using values of  $s$  lower than 0.3 does not alter the qualitative results much.

$$t^* \in \arg \max T = (tW) \text{Probability}(s^* \leq s)$$

The first order condition is

$$\frac{\partial T}{\partial t} = [1 - \Phi(s^*)] - t\phi(s^*) \frac{\partial s^*}{\partial t} \leq 0; \quad t \geq 0 \quad (5.2)$$

The second order condition is given by

$$\frac{\partial^2 T}{\partial t^2} = -\Psi \left[ 2\phi(s^*) - t\Phi''(s^*) \right].$$

Hence, a sufficient condition for the second order condition to hold is that the distribution function be (weakly) convex i.e.  $\Phi''(s^*) \geq 0$ . The comparative static properties of the optimal tax rate are summarized in the following proposition.

**Proposition 3** : *If  $\Phi''(s^*) \geq 0$  and  $\phi(s)$  is log concave then the optimal tax rate  $t^*$  is increasing in  $\lambda$ ,  $p$ ,  $\theta$  and  $\beta$ .*

Several commonly used density functions, such as the Gaussian, exponential, logistic, chi-squared, Laplace and the uniform are log concave; for a more complete list see Bagnoli and Bergstrom (1989). From Proposition 2, an increase in  $\lambda$ , and  $p$  reduces the proportion of the population that engages in tax evasion. At the optimal tax rate, an increase in the tax rate produces only second order effects on the total tax revenues, allowing an increase in the optimal tax rate. Using Proposition 2 an increase in  $\beta$  and in the loss aversion parameter,  $\theta$ , also reduce the equilibrium proportion of tax evaders and, hence, using similar reasoning to the previous case, allows for an increase in the optimal tax rate.

**Example 1** : *A particularly simple and tractable case of a log concave distribution is the uniform so that  $s \sim U \left[ \begin{matrix} s, \bar{s} \\ - \end{matrix} \right]$ . Defining  $\Delta s = \bar{s} - s$ , the first order condition 5.2 can be written as*

$$\frac{\partial T}{\partial t} = \left[ 1 - \frac{s^* - s}{\Delta s} \right] - \frac{t\Psi}{\Delta s} \leq 0; \quad t \geq 0. \quad (5.3)$$

For an interior solution, one can solve (5.3) to get the following simply formula for the optimal tax rate

$$t^* = \frac{\bar{s}}{2\Psi} \quad (5.4)$$

To calibrate the model, we use the values of  $\theta$  and  $\beta$  suggested by Tversky and Kahneman (1992) i.e.  $\theta = 2.25$ ,  $\beta = 0.88$ . For audit probability and the penalty rate we use conservative values suggested by Bernasconi (1998) i.e.  $p = 0.03$ ,  $\lambda = 2.0$ . Finally, for the domain of the uniform distribution we use  $s \sim U[0, 1.5]$ , which is likely to be a fairly conservative figure. Substituting these calibrated values in (5.4) one gets the optimal tax rate to be 33.33 percent which is not unrepresentative of Western democracies. What is reassuring about the optimal tax analysis is that a relatively conservative formulation generates reasonably realistic estimates of the optimal tax. It is also suggestive of the further use of prospect theory in a wider range of optimal tax models.

## 6. Conclusions

Why do people bother to pay taxes when, given relatively low audit probabilities and penalties, evasion should be extremely attractive to expected utility maximizers? This is puzzling, given “too much observed compliance” relative to the predictions of expected utility theory. This paper considers an alternative theoretical model that is based on cumulative prospect theory. Prospect theory characterizes individuals as loss averse with respect to some reference income. Furthermore, these individuals overweight small probabilities while underweighting large ones. Since prospect theory is derived from experimental evidence based on individual choice under generic situations of risk, it does not rely on purely technical restrictions to preferences, that are not based on independent empirical evidence and, hence, whose plausibility might be difficult to judge.

Results in the paper show that despite low audit probabilities and penalty rates, the magnitude of tax evasion predicted by prospect theory is consistent with the data. Individuals are also predicted to respond to an increase in the tax rate by increasing the amount evaded. This accords with intuition and evidence, but contrasts with the converse prediction made by expected utility theory. Finally, the optimal tax rates predicted by prospect theory, in the presence of tax evasion behavior, are consistent with actual tax rates.

Prospect theory was developed to explain actual choice under generic situations of risk. In applying prospect theory, we do not ‘fix’ the parameter values of the model to make it fit the evidence on tax evasion. Instead we use the parameters of human choice that are revealed from independent experimental evidence. Prospect theory is able to explain the tax evasion puzzles. Its predictions about the magnitudes of optimal income taxes in the presence of tax evasion are indicative of actual magnitudes. Hence, the paper concludes that the behavior of tax payers provides strong support for prospect theory.

## 7. Appendix

### Convex Indifference Curves Under Prospect Theory and alternative reference incomes

Suppose now that reference income equals  $\nu(1-t)W$  where  $\nu > 0$ ; the text assumes that  $\nu = 1$ . Define by  $I_A$  and  $I_{NA}$  respectively, the income of the taxpayer, relative to the reference point, when the taxpayer is and is not audited. It is easy to check that

$$I_A = W [(1-t)(1-\nu) - \lambda t - s] - D [\lambda t + s]$$

$$I_{NA} = W [1 - \nu(1-t)] - tD$$

In order to define the budget constraint, notice that at  $D = 0$ ,  $I_A = W [(1-t)(1-\nu) - \lambda t - s]$  and  $I_{NA} = W [1 - \nu(1-t)]$  while at  $D = W$ ,  $I_A = I_{NA} = W [(1-t)(1-\nu)]$ . From this it is easy to compute the slope of the budget constraint in  $I_A, I_{NA}$  space to be  $-\left[\lambda + \frac{s}{t}\right]$ , which is identical to the one in the text with  $\nu = 1$ , hence, movements of the reference point do not affect the slope of the budget constraint. Therefore, in terms of Figure 3.1, one proceeds as before but by defining a new origin (corresponding to a new point  $A$ ) with coordinates given by  $\{W(1-t)(1-\nu), W(1-t)(1-\nu)\}$ .

Also check that the slope of the indifference curve in  $I_A, I_{NA}$  space is now given by:

$$\frac{dI_A}{dI_{NA}} = \frac{-w(1-p)}{\theta w(p)} \left[ \frac{-I_A}{I_{NA}} \right]^{1-\beta} < 0$$

The sign follows because  $I_A$  is a loss relative to the reference point. Differentiating again, it is easy to show that  $\frac{d^2 I_A}{dI_{NA}^2} \geq 0$ . In terms of Figure 3.1, one now gets smooth, convex downward sloping indifference curves even in the prospect theory case. This should typically give an interior solution to the tax evasion decision compared to the bang-bang solution one gets under linear indifference curves.

#### *Proof of Proposition 2*

Using 3.8, and the definition of  $h(\cdot)$  in (3.7) and by defining  $t_c = s \left[ \left( \frac{w(1-p)}{\theta w(p)} \right)^{1/\beta} - \lambda \right]^{-1} \geq 0$  it follows that when  $t = t_c$ ,  $h = 0$ . When  $t < t_c$ ,  $h < 0$  and so  $D = W$ , whereas if  $t > t_c$ ,  $h > 0$  and so  $D = 0$ . The remaining part of the proposition follows by proceeding in a similar manner and defining  $\lambda_c = \left( \frac{w(1-p)}{\theta w(p)} \right)^{1/\beta} - \frac{s}{t}$  and  $s_c = t \left\{ \left( \frac{w(1-p)}{\theta w(p)} \right)^{1/\beta} - \lambda \right\}$ . ■

#### *Proof of Proposition 3*

Implicitly differentiate 5.2 successively with respect to  $\lambda$ ,  $p$ ,  $\theta$  and  $\beta$ . The generic derivative is

$$\frac{\partial t^*}{\partial x} = t^* \left( -\frac{\partial^2 T}{\partial t^2} \right)^{-1} \left[ \left( \frac{\phi(s^*)}{1 - \Phi(s^*)} \right)^2 \frac{\partial}{\partial s} \left[ \frac{\phi(s^*)}{1 - \Phi(s^*)} \right] + 1 \right] \left( -\frac{\partial \Psi}{\partial x} \right), \quad (7.1)$$

where  $x = \lambda, p, \theta$  and  $\beta$  and  $\Psi$  is defined in (5.1). The condition  $\Phi''(s^*) \geq 0$  ensures that the second order condition holds while log concavity of  $\phi(s)$  ensures that  $\frac{\partial}{\partial s} \left[ \frac{\phi(s^*)}{1 - \Phi(s^*)} \right] \geq 0$ , hence, the sign of  $\frac{\partial t^*}{\partial x} = -\text{sign}\left(\frac{\partial \Psi}{\partial x}\right)$ . For  $x = \lambda, p, \theta$  and  $\beta$  and using (5.1) check that the following hold.

$$\frac{\partial \Psi}{\partial \lambda} = -t^* \leq 0$$

$$\frac{\partial \Psi}{\partial \theta} = t^* \left( \frac{-1}{\beta} \right) \theta^{-\beta-1} [\exp((-\ln p)^\alpha - (-\ln(1-p))^\alpha)]^{\frac{1}{\beta}} \leq 0$$

$$\frac{\partial \Psi}{\partial \beta} = \left( \frac{-t^*}{\beta^2} \right) \frac{\log \left[ [\exp((-\ln p)^\alpha - (-\ln(1-p))^\alpha)]^{\frac{1}{\beta}} \right]}{[\exp((-\ln p)^\alpha - (-\ln(1-p))^\alpha)]^{\frac{-1}{\beta}}} \leq 0$$

$$\frac{\partial \Psi}{\partial p} = \left( \frac{-\alpha t^*}{\beta \theta} \right) \frac{\left[ \frac{(-\ln p)^{\alpha-1}}{p} + \frac{(-\ln(1-p))^{\alpha-1}}{(1-p)} \right]}{[\exp((-\ln p)^\alpha - (-\ln(1-p))^\alpha)]^{1-\frac{1}{\beta}}} \leq 0$$

The signing of these derivatives uses the fact that  $\exp((-\ln p)^\alpha - (-\ln(1-p))^\alpha) \geq 0$ , which is a necessary condition for  $s^* \in \text{int} \left[ \underline{s}, \bar{s} \right]$  (see 5.1); this follows for any of the parameter values used in this paper. The statement in the proposition now follows by using (7.1) in conjunction with the sign of the derivatives of  $\Psi$  with respect to the various parameters. ■

## 8. References

### References

- [1] al-Nowaihi, A., and Pyle, D. (2000) "Income tax evasion: A theoretical analysis" in MacDonald, A. and Pyle, D. *Illicit activity: The economics of crime and tax fraud* (Ashgate:Aldershot).
- [2] Andreoni, J., Erard, B., and Feinstein, J. (1998) "Tax Compliance" *Journal of Economic Literature*, 36: 818-860.
- [3] Allingham, M. G. and Sandmo, A. (1972) "Income tax evasion: A theoretical analysis" *Journal of Public Economics*, 1: 323-38.

- [4] Alm, J., McClelland, G. H., and Schulze, W. D. (1992) “Why do people pay taxes?” *Journal of Public Economics*, 48: 21-38.
- [5] Bagnoli, M., and Bergstrom, T. (1989) “Log-concave probability and its applications” Discussion Paper 89-23, University of Michigan.
- [6] Benartzi, S. and Thaler, R. H. (1995) “Myopic loss aversion and the equity premium puzzle” *Quarterly Journal of Economics*, 110: 73-92
- [7] Bernasconi, M., (1998) “Tax evasion and orders of risk aversion” *Journal of Public Economics*, 67: 123-34.
- [8] Bernasconi, M., and Zanardi, A. (2002) “Tax evasion, tax rates and reference dependence” mimeo.
- [9] Besley, T. and Coate, S. (1992) “Understanding welfare stigma: Taxpayer resentment and statistical discrimination” *Journal of Public Economics*, 48: 165-183.
- [10] Camerer, C. (2000) “Prospect theory in the wild: Evidence from the field” in Kahneman, D. and Tversky, A. *Choices, values and frames* (Cambridge: Cambridge University Press).
- [11] Camerer, C.F., and Ho, T. H. (1994) “Violations of the betweenness axiom and non-linearity in probability” *Journal of Risk and Uncertainty*, 8: 167-196.
- [12] Eide, E. (2001) “Rank dependent expected utility models of tax evasion” mimeo, University of Oslo.
- [13] Gordon, J. P. F. (1989) “Individual morality and reputation costs as deterrents to tax evasion” *European Economic Review*, 33: 797-805.
- [14] Kahneman, D. and Tversky, A. (1979) “Prospect theory: An analysis of decision under risk” *Econometrica*, 47: 263-91.
- [15] Mankiw, N. G. and Zeldes, S. (1991) “The consumption of stockholders and non stockholders” *Journal of Financial Economics*, 29: 97-112.
- [16] Myles, G. (1995) *Public Economics* (Cambridge: Cambridge University Press).
- [17] Prelec, D. (2000) “Compound invariant weighting functions in prospect theory” in Kahneman, D. and Tversky, A. *Choices, values and frames* (Cambridge: Cambridge University Press)
- [18] Prelec, D. (1998) “The probability weighting function” *Econometrica*, 60: 497-528.

- [19] Pudney, S., Pyle, D. and Saruc, T. (2000) "Income tax evasion: An experimental approach" in MacDonal, A. and Pyle, D. *Illicit activity: The economics of crime and tax fraud* (Ashgate:Aldershot).
- [20] Segal, U., and Spivak, A. (1990) "First order versus second order risk aversion" *Journal of Economic Theory*, 51: 111-25.
- [21] Starmer, C. (2000) "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk" *Journal of Economic Literature*, 38: 332-82.
- [22] Skinner, J., and Slemrod, J. (1985) "An economic perspective on tax evasion" *National Tax Journal*, 38: 345-53.
- [23] Tversky, A. and Kahneman, D. (1992) "Advances in prospect theory: Cumulative representation of uncertainty" *Journal of Risk and Uncertainty*, 5: 297-323.
- [24] Yaniv, G., (1999) "Tax compliance and advance tax payments: A prospect theory analysis" *National Tax Journal*, 52: 753-64.
- [25] Yitzhaki, S. (1974) "A note on income tax evasion: A theoretical analysis" *Journal of Public Economics*, 3: 201-202.