

DEPARTMENT OF ECONOMICS

**OPTIMAL INCOME TAXATION IN THE
PRESENCE OF TAX EVASION: EXPECTED
UTILITY VERSUS PROSPECT THEORY**

**Sanjit Dhami, University of Leicester, UK
Ali al-Nowaihi, University of Leicester, UK**

**Working Paper No. 07/10
September 2007**

Optimal income taxation in the presence of tax evasion: Expected utility versus prospect theory.

Sanjit Dhami* Ali al-Nowaihi†

27 September 2007

Abstract

The predictions of expected utility theory (EUT) applied to tax evasion are flawed on two counts: (i) They are quantitatively in error by huge orders of magnitude. (ii) Higher taxation is predicted to lower evasion, which is at variance with the evidence. An emerging literature in behavioral economics, most notably based on prospect theory (PT), has shown that behavioral economics is much better at explaining tax evasion. We extend this literature to incorporate issues of optimal taxation. As a benchmark for a successful theory, we require that it should explain, jointly, the facts on the tax rate, tax gap and the level of government expenditure. We find that when taxpayers use EUT (respectively, PT) and the optimal tax is derived from a social welfare function that also uses EUT (respectively, PT), then, the calibration results are completely at odds with the facts. However, when taxpayers use PT but the social welfare function uses standard EUT, there is a very close match between the predictions and the facts. This has important implications for context dependent preferences but also for the newly emerging literature on liberalism versus paternalism in behavioral economics.

Keywords: Prospect theory, Expected utility theory, Tax evasion, Optimal taxation, Normative versus positive economics, Context dependent preferences, Liberalism, Paternalism.

JEL Classification: D81 (Criteria for Decision Making Under Risk and Uncertainty), 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.

†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.

“Even when an agent is perfectly rational in the sense that he systematically maximizes some function, it is not at all obvious that the utility function which explains his behavior should be inserted into the welfare considerations ... More generally, once psychological effects enter into the calculus, there is no escape from separating welfare and behavior.” Rubinstein (2005).

“paternalism [is the] power and authority one person or institution exercises over another to confer benefits or prevent harm for the latter regardless of the latter’s informed consent ... paternalism ... increases the potential for the abuse of state power, arbitrary discrimination, tyranny, and civil strife ... liberalism ... provides the only way for the proponents of conflicting ways of life to live together ... liberalism [is] the only viable basis for peaceful coexistence in culturally and religiously plural societies.” The Oxford Companion to Philosophy (2005, pp 515, 648).

1. Introduction

Issues of tax evasion are extremely important for all countries. Losses to society from tax evasion are huge. For the USA, for example, based on the most recent data, the *tax gap*¹ is of the order of \$300 billion per year (Slemrod, 2007).² An important feature of the existing analysis of tax evasion is that it has largely been carried out in an expected utility theory (EUT) framework.

Recent research points to several serious problems in using an EUT approach to tax evasion. Dhami and al-Nowaihi (2007) apply Kahneman and Tversky’s (1979) prospect theory³ (PT) to the tax evasion decision facing a taxpayer. They show that while EUT gives the correct qualitative results for the effects of the probability of detection and the penalty rate, there are several problems. First, EUT makes the prediction that under reasonable attitudes to risk, namely, non-increasing absolute risk aversion, the taxpayer evades less as the tax rate goes up. The implication is that tax evasion will be at a minimum when the tax rate is 100 percent. This result, due to Yitzhaki (1974), is contradicted by the bulk of empirical evidence. Second, at existing penalty rates and detection probabilities, the quantitative predictions of EUT on the extent of tax evasion are wrong by a factor of about 100. On the other hand, PT gives the correct quantitative and qualitative results.⁴

¹The *tax gap* is the difference between the amount owed in taxes and the amount actually collected by the tax authorities. Income tax accounts for about two thirds of this gap; see Slemrod (2007).

²To this has to be added the cost of enforcement, the cost of misallocation of resources due to workers and firms diverting their efforts to less productive but easier to evade activities and the cost due to distorted prices, in particular between privately and publicly provided goods and services.

³The standard references for prospect theory are Kahneman and Tversky (1979) (which, incidentally, is the most cited Econometrica paper) and Tversky and Kahneman (1992). Subsection 4 below gives a self-contained exposition of prospect theory.

⁴But these are not the only problems that PT can rectify in the context of tax evasion. Empirical and

Given the magnitudes involved, the misleading welfare consequences of applying EUT to an analysis of tax evasion are potentially very large.

Dhami and al-Nowaihi (2007), however, treat the tax rate as exogenous. In this paper, we extend the analysis of Dhami and al-Nowaihi (2007) by asking *what should the optimal income tax be, when taxpayers use PT to make their tax evasion decision?* This is a substantially more difficult question because the appropriate welfare criteria under behavioral economics is, as yet, an unsettled area.

1.1. A brief description of the model

Our framework of analysis is as follows. We consider a model where the government levies taxes to finance public provision of goods and services. Individuals can choose to evade a fraction of their income. The government audits a fraction of the tax returns. If a taxpayer is caught evading, he pays back owed tax plus a penalty. Individuals gain utility from both private and public consumption. The government chooses the optimal tax rate, given society's preference between private and public expenditure, and taking the subsequent tax evasion behavior of taxpayers into account. In this simple framework, we assess the relative success of EUT and PT. We find that PT far outperforms EUT, and also highlights the importance of recognizing that preferences are context dependent.

1.2. Brief literature review

The literature on endogenous evasion and optimal taxation is fairly limited and, without exception, uses EUT. There are three main strands of the literature. In the first strand, associated with Cremer and Gahvari (1993) and Boadway, Marchand and Pestieau (1994), the problem is to find the optimal taxes set by a benevolent planner in the presence of *tax avoidance*, rather than *tax evasion*. So, by incurring a cost, the taxpayer can ensure that evasion is never discovered by the tax authorities. This in itself is an interesting problem and allows one to make a case for commodity taxation. Because labour income can be avoided while commodity income cannot be, hence, it is efficient to have some commodity taxes⁵. However, this leaves open the relation between tax evasion and optimal taxation.

The second strand is exemplified in the work of Cremer and Gahvari (1996), Marhuenda and Ortuno-Ortin (1997) and Chandar and Wilde (1998). Here the approach is to choose, simultaneously, the optimal tax and the enforcement structure in order to induce truthful

experimental evidence show that *obligatory advance tax payments* reduce tax evasion, a fact that can be explained by PT but not by EUT; see Elffers and Helsing (1997) and Yaniv (1999).

⁵In the absence of tax evasion it is known, from the results of Atkinson and Stiglitz (1972, 1976), that commodity taxes are redundant if income taxes are available. However, if tax bases are measured with error, then a case can be made for commodity taxes even in the absence of tax evasion; see Dhami and al-Nowaihi (2006).

reporting of income. In equilibrium, there is no tax evasion. However, if eliminating evasion completely is not possible, then the question of designing an optimal tax structure remains unanswered.

The third strand is given by the recent work of Richter and Boadway (2006) to which our work is most closely connected. They explicitly model the tax evasion decision in the standard Allingham and Sandmo (1972) model. They consider a representative household, hence, the focus is entirely on efficiency issues. The central question, as in the first strand, is the optimal mix of income and consumption taxes when the ease of evading different types of taxes is different. The welfare objective of the government is to maximize its total revenues (arising from income taxes, consumption taxes and penalties on those caught evading) subject to a taxpayer participation constraint. The main disadvantage of the income tax is that, by inviting tax evasion, it exposes the risk-averse taxpayer to income risk.⁶ While the consumption tax cannot be evaded, its main disadvantage is that it distorts relative prices of goods. Hence, optimal taxation arises from a tradeoff between imposing income risk on the taxpayer and creating tax distortions.

In the context of our paper, the main criticisms of the existing literature are as follows. First, the behavioral approach to tax evasion shows that an EUT based analysis of tax evasion is seriously misleading. However, the existing literature is based entirely on the EUT framework. Second, tax evasion is a very real phenomenon in both developing and developed countries. The mechanism design approach to evasion, which essentially searches for optimal tax/penalty schemes to completely eliminate evasion, does not seem suitable. Third, the cost of risk to the taxpayer from the risky activity of tax evasion needs to be explicitly modelled. Fourth, one needs to explore welfare criteria other than revenue maximization in designing optimal tax schemes.

1.3. Welfare analysis under behavioral economics

Under uncertainty, when decision makers have expected utility preferences, the Pareto frontier can be found, subject to the standard regularity conditions, by a benevolent utilitarian planner who maximizes a weighted sum of the individuals' expected utilities.

Suppose now that individuals have *prospect theory preferences* (which we discuss in detail in section 4). Under these preferences, decision makers (1) overweight small probabilities, (2) evaluate gains and losses relative to a reference point, (3) are loss-averse and (4) have distinct risk preferences depending on whether they are in the domain of gains or losses, relative to the reference point.⁷

⁶This exposure to risk is completely ignored by the first strand of the literature, which assumes a constant cost of such risky activity.

⁷Behavioral economics differentiates between several types of utility, in particular between *decision utility* and *experienced utility*. PT, being a decision theory like EUT, considers decision utility; see, for

Should a social planner respect the PT preferences of an individual? Or should the planner disregard these preferences and evaluate the well being of society on the basis of expected utility? These questions take us into relatively un-chartered territory in economics. There is no clear consensus on the approach to be taken, but we illustrate some well known views on this matter, below.

Tversky and Kahneman (1986, abstract) state that “no theory of choice can be both normatively adequate and descriptively correct”. Their argument is as follows. Since *invariance* and *dominance* are always obeyed when their application is transparent, they should be essential features of any normative theory. However, because of bounded rationality, they are often violated when application is not transparent. A descriptively adequate theory must take account of this.

The modern literature on merit goods has stressed the bounded rationality of decision makers. Should policy makers respect boundedly rational decisions or try to alter them? This is the theme of recent work by Camerer et. al. (2003). These authors advocate the case for *asymmetric paternalism*. Essentially this is an attempt to steer boundedly rational people in the direction of avoiding costly mistakes but, at the same time, distorting the decisions of rational people as little as possible. In the words of the authors: “And a variety of researchers have shown that people exhibit systematic mis-predictions about the costs and benefits of choices—for example, the degree of loss aversion exhibited in people’s choices seems inconsistent with their actual experiences of gains and losses. It is such errors—apparent violations of rationality—that can justify the need for paternalistic policies to help people make better decisions and come closer to behaving in their own best interest.”⁸

Several other forms of paternalism are advocated in the literature. Benjamin and Laibson (2003) introduce the concept of *benign paternalism*. The idea here is to encourage the individual to undertake socially desirable actions, without violating individual liberty to make the decision. The idea is made operational by introducing small hurdles in the way of harmful individuals choices in order to shepherd them in a better direction.⁹ Jolls et al (1998) coin the term *anti-antipaternalism* which rejects the idea of pure libertarianism. O’Donoghue and Rabin (2003) advocate the idea of *optimal paternalism* which advocates taking account of all costs and benefits of paternalism. They reject the view

example, Kahneman and Tversky (2000, section 9). Thus all utilities in this paper are decision utilities.

⁸Camerer et. al. (2003) illustrate the usefulness of their approach for boundedly rational individuals in several contexts such as the following. Setting default options which encourage savings behavior and protect insurance rights, framing of contracts to include seemingly irrelevant information, disclosure issues, etc.

⁹For instance, gamblers are asked to choose up-front a level of liquidity which cannot be subsequently exceeded when they might be in a tempted state. Gamblers without subsequent self control problems will not need to be disciplined so they have no problems with setting liquidity limits, while gamblers with subsequent self control problems will clearly benefit from these up-front limits.

that interventions should be minimal. In their words: “In some instances, even seemingly large deviations from the policy that is optimal for fully rational economic agents would not cause severe harm to those agents. In such cases, even a small probability of people making errors can have dramatic effects for optimal policy.”

1.4. Liberalism, paternalism and context dependence of preferences

In section 1.3 we have illustrated the increasing appeal of paternalism (as opposed to liberalism) in conducting welfare analysis when individuals have bounded rationality. However, *non-paternalism*¹⁰ is a fundamental principle of liberalism: individuals are the best judges of their own welfare. It is also a fundamental assumption of social choice theory, welfare economics and mechanism design.¹¹ In its weakest and most general form, non-paternalism is expressed by the *Bergson-Samuelson* social welfare function:

$$W(\mathbf{x}) = F(u_1(\mathbf{x}), u_2(\mathbf{x}), \dots, u_I(\mathbf{x})) \quad (\text{Bergson-Samuelson}), \quad (1.1)$$

where \mathbf{x} is a social state, I is the number of individuals in society and $u_i(\mathbf{x})$ is the utility of individual i as seen by that individual. A special case of (1.1) is that of *constant elasticity*:

$$W(\mathbf{x}) = \frac{1}{1-\rho} \sum_{i=1}^I \beta_i [u_i(\mathbf{x})]^{1-\rho}, \quad \rho \neq 1, \beta_i \geq 1, u_i(\mathbf{x}) > 0 \quad (\text{constant elasticity}), \quad (1.2)$$

where the government exhibits a degree of paternalism in that (1) it may give different weights to different individuals and (2) it may exhibit a degree of inequality aversion (captured by ρ) different from that exhibited by an individual (as captured by $u_i(\mathbf{x})$). Further specialization of (1.2) gives the two forms we shall use in this paper¹²:

$$W(\mathbf{x}) = u_i(\mathbf{x}) \quad (\text{representative agent}), \quad (1.3)$$

$$W(\mathbf{x}) = \sum_{i=1}^I u_i(\mathbf{x}) \quad (\text{utilitarian or Benthamite}), \quad (1.4)$$

Form (1.4) may be viewed as the most liberal of the above, as the government gives equal weight to all individuals and does not modify any of the individual utility functions. A number of axiomatizations lead to (1.4).¹³

¹⁰Also known as *welfarism* or *individualism*.

¹¹See, for example, chapters 21, 22 and 23 in Mas-Colell, Whinston and Green (1995).

¹²We prefer the term *representative agent* to *dictator*. First, individual i may be different from the decision maker. Second, even if individual i is the decision maker, he/she might be selected by a democratic process, say an election.

¹³For example, Harsanyi (1953, 1955, 1975) derived (1.4) using expected utility theory and a ‘behind the veil of ignorance’ argument. D’Aspremont and Gevers (1977) derived (1.4) from cardinal measurability

We shall argue that the evidence on tax evasion is best explained by the following combination. We use PT to model the tax evasion decision and derive the government budget constraint. The government then chooses the tax rate that maximizes social welfare. Although the government recognizes that the tax evasion decision is best described by PT, when it chooses the tax rate it assumes that society's preference over private and public provision, the $u_i(\mathbf{x})$ in (1.3) or (1.4), is given by standard utility theory. We could defend this methodology in either of two ways. The first is more traditional while the latter relies on emerging empirical evidence. Our results *do not* hinge on which interpretation we adopt. Readers of different persuasions will have a preference for one or the other.

1. The government, when choosing the tax rate, behaves paternalistically, i.e., the government assumes that it has better knowledge of the true welfare of individuals than the individuals themselves.

2. Paternalism can have negative connotations (see our second quote at the beginning of the paper). Hence, our own preference is for an alternative view that draws inspiration from a very large body of empirical evidence that has been generated in behavioral economics. In particular, the evidence suggests, very clearly, that individuals do not have a complete preference ordering over all states and that preferences are heavily *context dependent*. Alternative contexts can arise, for instance, from the framing of choices.¹⁴ Context dependent preferences do not go away once professionals are presented with choices that they must make on a regular basis.¹⁵ The mental accounting literature pioneered by Richard Thaler is also suggestive of context dependent preferences.¹⁶ Individuals, when making a private consumption decision might act so as to maximize their selfish interest. But in a separate role (context), say, as part of the government, as a school governor or as a voter, could act so as to maximize some notion of public well being. Many examples can be given.

1. Individuals might send their own children to private schools (self interest) but could at the same time vote for more funding to government run schools in local or national

with unit comparability, non-paternalism, the strong Pareto principle and anonymity (or symmetry). Maskin (1978) derived (1.4) from cardinal measurability, non-paternalism, the strong Pareto principle, anonymity, separability of indifferent households, continuity and strong equity. Other axiomatizations lead to other special forms of (1.1). For a review see, for examples, chapters 5 and 6 of Boadway and Bruce (1984).

¹⁴The following is just one example out of hundreds described in Kahneman and Tversky (2000). It is problems 9 and 10 from Quattrone and Tversky (1988). In a survey, 64% of respondents thought that an increase in inflation from 12% to 17% was acceptable if it lead to a reduction in unemployment from 10% to 5%. However, only 46% of the respondents thought that exactly the same increase in inflation (from 12% to 17%) was acceptable if it increased employment from 90% to 95%.

¹⁵In a well known example, Kahneman and Tversky find, in the context of medical decisions, that the choice between various programs depends on whether the choices are posed in terms of *lives saved* or *lives lost* (see page 5 in Kahneman and Tversky (2000)).

¹⁶See Part 4 of the book by Kahneman and Tversky (2000).

elections (public interest).

2. In making their labor supply decision, individuals might tradeoff personal utility from consumption with disutility from labor supply. But the same individual, in his/ her voting choice, might vote for redistributive policies that exhibit public concern (and not purely private self interest).
3. An individual, when buying an air ticket, might also buy travel insurance, thus exhibiting risk averse behavior. But, the same individual, when he reaches his holiday destination, may visit a gambling casino and exhibit risk loving behavior there.

Using this view, we can assume that individuals, when taking their tax evasion decision, exhibit behavior described by PT. But when individuals express their preferences over private and public provision (through, say, surveys, referenda, elections, etc.), these preferences are described by standard utility theory. This could be formalised by allowing the utility function, $U(z, \cdot)$, to depend on a *context variable*, z . So in our case, for example, $z = 1$ when an individual is deciding how much tax to evade but $z = 2$ when the same individual is voting for a tax system.

1.5. Results and schematic outline

The results are as follows. Under both EUT and PT, taxpayers evade less as the audit probability and the penalty rate increase. Under PT, additionally, one gets the plausible result that taxpayers evade more if the tax rate increases. Using EUT to model both tax evasion and government behavior is unable to jointly account for the evidence on actual tax gaps and government expenditure. Given actual penalty rates and audit probabilities, if consumer preferences over private versus public consumption yield observed government expenditures, then EUT predicts far too big a tax gap. On the other hand, if consumer preferences over private and public consumption yield observed tax gaps, then EUT predicts far too much government expenditure. Using PT to model both tax evasion and government behavior gives economically absurd results. By contrast, using PT to model the tax evasion decision and standard utility theory to explain government behavior, we have no difficulty reconciling observed tax gaps with observed government expenditures, at plausible tax rates.

Section 2 describes the basic model. Sections 3 and 4 consider the tax evasion decision on the basis of EUT and PT, respectively. Sections 5 and 6 derive the resulting optimal tax rates under EUT and PT, respectively. Section 7 compares the success of PT in explaining tax evasion with that of EUT. Finally, section 8 summarizes and concludes.

2. The Model

We consider an economy consisting of a continuum of consumer-taxpayers located on the unit square $\Omega = [0, 1] \times [0, 1]$. Pretax income is exogenous and is given by the density function, $Y(x, s) \geq 0$, $x \in [0, 1]$, $s \in [0, 1]$, where s (explained in more detail below) captures the *stigma* faced by a tax evader when caught and x is purely a label that helps locate an individual taxpayer.¹⁷ The government levies tax at the constant rate $t \in [0, 1]$ on declared income and uses its tax revenue to finance the public provision of goods and services whose aggregate monetary value is G .¹⁸

The consumer-taxpayer located at (x, s) declares income $D(t, G, x, s) dx ds$, $D(t, G, x, s) \in [0, Y(x, s)]$. Thus $D(t, G, x, s)$ is the density of declared income. Subsequent to the filing of tax returns, an exogenous fraction $p \in (0, 1)$ of the taxpayers are audited, and the audit reveals the true taxable income. If caught, the dishonest taxpayer must pay the outstanding tax liabilities $t[Y(x, s) - D(t, G, x, s)] dx ds$ and a penalty proportional to unpaid taxes, $\lambda t[Y(x, s) - D(t, G, x, s)] dx ds$, where $\lambda > 0$ is the constant penalty rate. The density of tax revenue is then

$$T(t, G, x, s) = tD(t, G, x, s) + p(1 + \lambda)t[Y(x, s) - D(t, G, x, s)]. \quad (2.1)$$

Clearly, D and T depend on p and λ as well as t , G , x and s .

2.1. Sequence of moves

The sequence of moves is as follows.

1. The tax authority announces the tax rate, t , the audit probability, p , the penalty rate, λ , and the monetary value of the provision of publicly provided goods/services, G .
2. Taxpayers make the decision to either report full income or evade a fraction of it, given t , p , λ and G .
3. 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.

¹⁷In the discrete case, x could denote the individual's *social security number* or his *national insurance number* which uniquely locates the individual.

¹⁸Since Ω is of unit area, G can be interpreted as either the aggregate or the per capita monetary value of all publicly provided goods and services.

2.2. Exogenous and endogenous variables

The exogenous variables of the model are the probability of an audit, p (which, here, is the same as the probability of detection), the penalty rate, λ , and the density function of income, Y . When considering tax evasion under prospect theory, we shall introduce and explain the further exogenous variables α , β and θ . The endogenous variables of the model are the tax rate, t , the density function of declared income, D , and the tax revenue density function, T . When the consumer-taxpayer makes his tax evasion decision, he takes as exogenous the model parameters, his location, (x, s) , the tax rate, t , and the level of public provision, G .

Notation: To simplify notation, we shall suppress reference to the model parameters and drop the infinitesimal quantities, dx and ds , when such omission is not likely to lead to confusion.

2.3. Government tax revenue

We assume that Y is integrable¹⁹. It will follow from optimizing behavior of consumer-taxpayers that D is integrable. Hence, T is also integrable. Let $S \subset \Omega$ be measurable²⁰. Then the aggregate pretax income of all consumer-taxpayers in S is given by

$$\bar{Y}(S) = \int \int_{(x,s) \in S} Y(x, s) ds dx, \quad (2.2)$$

and the tax revenue collected from them is

$$\begin{aligned} \bar{T}(t, G, S) &= \int \int_{(x,s) \in S} t D(t, G, x, s) dx ds \\ &+ p(1 + \lambda) \int \int_{(x,s) \in S} t [Y(x, s) - D(t, G, x, s)] dx ds. \end{aligned} \quad (2.3)$$

In particular, let

$$\bar{Y} = \bar{Y}(\Omega) = \int_{x=0}^1 \int_{s=0}^1 Y(x, s) ds dx, \quad (2.4)$$

and

$$\begin{aligned} \bar{T}(t, G) &= \bar{T}(t, G, \Omega) = \int_{x=0}^1 \int_{s=0}^1 t D(t, G, x, s) dx ds \\ &+ p(1 + \lambda) \int_{x=0}^1 \int_{s=0}^1 t [Y(x, s) - D(t, G, x, s)] dx ds, \end{aligned} \quad (2.5)$$

¹⁹Integrability can be interpreted either in the sense of Riemann or Lebesgue, it does not matter for our purposes which interpretation is chosen. However, the Lebesgue integral is more general than that of Riemann, in that it is defined for wider classes of functions and domains of integration. It is also associated with powerful convergence theorems. A particularly clear introduction is Bartle (1966).

²⁰Measurable in the sense of Borel or Lebesgue. Any set of interest to us will be measurable in both senses.

then \bar{Y} is both total income and average income. Likewise, $\bar{T}(t, G)$ is both total and average tax revenue. We assume that $\bar{Y} > 0$.

Note that if S is of measure zero then $\bar{Y}(S) = 0$ and, hence, also $\bar{T}(t, G, S) = 0$. In particular, the tax revenue collected from any single individual, (x, s) , is zero: $\bar{T}(t, G, \{(x, s)\}) = 0$. This ensures that we can consistently assume that when a taxpayer decides how much income to declare, he can take public provision, G , as given.²¹

Thus the government's tax revenue, $\bar{T}(t, G)$, comes from three sources:

1. taxes on declared income, $\int_{x=0}^1 \int_{s=0}^1 t D(t, G, x, s) dx ds$,
2. taxes recovered from those caught evading, $p \int_{x=0}^1 \int_{s=0}^1 t [Y(x, s) - D(t, G, x, s)] dx ds$,
and
3. fines from those caught evading, $p\lambda \int_{x=0}^1 \int_{s=0}^1 t [Y(x, s) - D(t, G, x, s)] dx ds$.

From (2.4) and (2.5), we see that total tax revenue can be written in the slightly simpler form:

$$\bar{T}(t, G) = tp(1 + \lambda) \bar{Y} + t[1 - p(1 + \lambda)] \int_{x=0}^1 \int_{s=0}^1 D(t, G, x, s) dx ds. \quad (2.6)$$

It will be useful to distinguish between those who are unable to evade tax, and those who can evade but choose not to. Therefore, assume taxes are deducted at source for a fraction, ω , of the population, where

$$0 \leq \omega \leq 1, \quad (2.7)$$

so, these taxpayers cannot evade (if $\omega = 1$, then nobody can evade).

We make the following two simplifying assumptions.

A1. *The opportunity to evade taxes does not depend on the taxpayer's income:* Thus, $Y(x, s)$ does not depend on ω . Hence, the tax revenue collected from the consumer-taxpayers in $S \subset \Omega$ is given by

$$\begin{aligned} \bar{T}(t, G, S) = & \omega t \int \int_{(x,s) \in S} Y dx ds \\ & + (1 - \omega) t \left[\int \int_{(x,s) \in S} D dx ds + p(1 + \lambda) \int \int_{(x,s) \in S} [Y - D] dx ds \right], \end{aligned} \quad (2.8)$$

²¹To be absolutely clear, $T(t, G, x, s)$ is the *density* of tax revenue at (x, s) . If $t > 0$ and $D(t, G, x, s) > 0$, then $T(t, G, x, s) > 0$. However, the *tax revenue* collected from (x, s) , $\bar{T}(t, G, \{(x, s)\})$, is zero. A useful intuitive picture to retain in mind is $\bar{T}(t, G, \{(x, s)\}) = T(t, G, x, s) dx dy$, where $T(t, G, x, s) > 0$ but $\bar{T}(t, G, \{(x, s)\}) = 0$ because the 'infinitesimal rectangle', $dx dy$, has zero area. Such ideas can be made rigorous using Non-standard analysis, which was invented for such purpose; see, for example, Loeb and Wolff (2000).

which simplifies to

$$\begin{aligned}\bar{T}(t, G, S) &= [\omega + p(1 + \lambda)(1 - \omega)]t \int \int_{(x,s) \in S} Y dx ds \\ &\quad + (1 - \omega)[1 - p(1 + \lambda)]t \int \int_{(x,s) \in S} D dx ds.\end{aligned}\tag{2.9}$$

Total tax revenue becomes

$$\begin{aligned}\bar{T}(t, G) &= \bar{T}(t, G, \Omega) \\ &= [\omega + p(1 + \lambda)(1 - \omega)]\bar{Y}t + (1 - \omega)[1 - p(1 + \lambda)]t \int_{x=0}^1 \int_{s=0}^1 D dx ds.\end{aligned}\tag{2.10}$$

A2. Stigma is unrelated to income: Although different people suffer different rates of stigma, we do not know of any strong evidence that this is related to income, e.g., we do not know of any strong evidence that rich persons suffer higher, or lower, rates of stigma than poor persons, on account of their income.²² Our second simplifying assumption is, therefore, that the income density function, $Y(x, s)$, is independent of s . From (2.4) we get

$$\bar{Y} = \int_{x=0}^1 \int_{s=0}^1 Y ds dx = \int_{x=0}^1 Y \left(\int_{s=0}^1 ds \right) dx = \int_{x=0}^1 Y dx.\tag{2.11}$$

The independence of income, stigma and the ability to evade, will give rise to simple expressions for total tax revenue and aggregate utility.

2.4. Behavior of government and consumer-taxpayers

The tax authority moves first, making an announcement of the tax rate, t , the audit probability, p , the penalty rate, λ , and the total monetary value of all publicly provided goods and services, G . Given t , p , λ and G , the taxpayer then makes the decision to either report full income ($D = Y$) or evade a fraction of it ($D < Y$). Let Y_{NC} be the after tax income of the taxpayer if he is not caught, then

$$Y_{NC} = Y - tD, \text{ with probability } 1 - p.\tag{2.12}$$

If evasion is discovered, the taxpayer also suffers some stigma, whose monetary value is $s(Y - D)$, where s is the stigma rate on evaded income, $s \in [0, 1]$. As in Gordon (1989)

²²See Slemrod (2007, p30) for a review of this.

and Besley and Coate (1992), such stigma enters linearly, as a monetary equivalent, into the payoff in that state of the world²³. His after-tax income is then Y_C , given by

$$Y_C = (1 - t)Y - (s + \lambda t)(Y - D), \text{ with probability } p. \quad (2.13)$$

The government spends the total tax revenue, $\bar{T}(t, G)$, given by (2.6), on providing goods and services under the balanced budget constraint,

$$\bar{T}(t, G) = G. \quad (2.14)$$

Consumers derive utility from private consumption and the publicly provided goods and services. However, when making the decision on how much income to declare, a consumer takes the publicly provided goods and services as given. Thus we have a free-rider problem: each consumer derives utility from public provision, but hopes others will pay for it. The government chooses the tax rate, t , so as to maximize social welfare, taking into account the utility individuals derive from private and public consumption and the effect of the tax rate on tax evasion and, hence, on tax revenue.

We shall consider five regimes, summarized in Table-I below. We use the shorthand notation CARA for constant absolute risk aversion and log for logarithmic utility.

Table-I: Classification of regimes

Regime	Taxpayer preferences	Government preferences
EUT	EU: complete	EU, representative consumer, separable, log
PT1	PT: context dependent	EU, representative consumer, separable, log
PT2	PT: context dependent	EU, representative consumer, non-separable
PT3	PT: context dependent	EU, utilitarian, separable, CARA
PT4	PT: complete	PT, utilitarian

We now describe the five regimes more fully. The description will rely heavily on the arguments developed in the introduction; see in particular section 1.4.

1. *Regime EUT*: The second row of Table-I describes regime EUT. In this regime, consumer preferences over private provision, C , and public provision, G , are given by the tractable form:

$$U(C, G) = \mu \ln G + (1 - \mu) \ln C, \quad 0 < \mu < 1, \quad G > 0, \quad C > 0, \quad (2.15)$$

Expected utility is then

$$EU = \mu \ln G + (1 - \mu) [p \ln Y_C + (1 - p) \ln Y_{NC}], \quad (2.16)$$

²³Although 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 reputational costs that impinge on current and future earnings. For a more detailed discussion of stigma in the context of tax evasion as well as a more general formulation, see Dharami and al-Nowaihi (2007).

where Y_C and Y_{NC} are given by (2.13) and (2.12), respectively.

Maximizing (2.16) gives the density function, D , of declared income and, hence, government tax revenue $\bar{T}(t, G)$. Thus, the consumer-taxpayer has a complete transitive set of preferences: the same utility function describes both tax evasion and preference between private and public provision. The government uses the same utility function (2.15) to maximize the utility of a representative consumer with average income, \bar{Y} , who does not evade. In other words, we do not need to invoke context dependence or paternalism to justify this regime. The details are given in section 3, below. This regime turns out to perform poorly in jointly explaining the tax rate, the tax gap and government expenditure.

2. *Regime PT1*: The third row of Table-I describes regime PT1. In this regime, consumer preferences over private and public provision are given by the same utility function (2.15) as in regime EUT. However, when making the tax evasion decision, consumer preferences are not given by expected utility (2.16) but by prospect theory (PT), as described in section 4, below. Thus, the consumer-taxpayer has *context-dependent preferences* (see the introduction): the utility function which describes tax evasion behavior is *different* from that which gives preferences between private and public provision. The government uses the same utility function (2.15), as in regime EUT, to maximize the utility of a representative consumer with average income, \bar{Y} , who does not evade. The details are given in section 4, below.²⁴ As discussed in the introduction, one could invoke paternalism or incomplete preferences to justify this approach. In contrast to regime EUT, regime PT1 successfully explains the tax rate, the tax gap and the level of government expenditure.
3. *Regime PT2*: The utility function (2.15) is additively separable over private and public provision. However, in actual practice, there would seem to be strong complementarities between the two. For example, utility derived from private car ownership heavily depends on the quality of publicly provided roads. We would like to investigate the consequences of recognizing this complementarity. The fourth row of Table-I describes regime PT2. In this regime, consumer preferences over private and public provision are given by the utility function:

$$U(C, G) = F\left(C - \frac{B}{G}\right), \quad F' > 0, \quad F'' < 0, \quad G > 0, \quad 0 < \frac{B}{G} < C, \quad (2.17)$$

For example, $F\left(C - \frac{B}{G}\right) = \ln\left(C - \frac{B}{G}\right)$. An interpretation of this model is that public provision, G , has value only insofar as it facilitates private consumption, C . When

²⁴How the government discovers that (2.15) correctly describes a consumer's preference between private and public provision, is not modelled in this paper (it could be through, e.g., surveys, elections, referenda, observing relevant market behavior, etc.).

making the tax evasion decision, consumer preferences are given by prospect theory, as in regime PT1. Thus, the consumer-taxpayer, again, exhibits context-dependent preferences. The government uses the utility function (2.17) to maximize the utility of a representative consumer with average income, \bar{Y} , who does not evade. The details are given in section 4, below. The results are very similar to that of regime PT2.

4. *Regime PT3*: The fifth row of Table-I describes regime PT3. In this regime consumer preferences over private and public provision are given by the utility function:

$$U(C, G) = \begin{cases} \mu \frac{G^{1-\gamma}}{1-\gamma} + (1-\mu) \frac{C^{1-\gamma}}{1-\gamma}; & \gamma \neq 1 \\ \mu \ln G + (1-\mu) \ln C; & \gamma = 1 \end{cases} \quad (2.18)$$

where $0 < \mu < 1$, $G > 0$, $C > 0$. When making the tax evasion decision, consumer preferences are given by prospect theory, as in regimes PT1 and PT2. Thus, the consumer-taxpayer, again, exhibits context-dependent preferences. The government uses the utility function (2.18) to maximize the *sum* (in the form of an integral) of the utilities of all the consumers in the state in which they do *not* evade. The details are given in section 4, below. Note that regime PT3 differs from regimes EUT and PT1 in two ways: the coefficient of relative risk aversion, γ , can now take any value (not just 1) and the government now maximizes the sum of all utilities, not that of a representative consumer. Yet the result is very close to those of regimes PT1 and PT2 and regime PT3 successfully explains the tax rate, tax gap and government expenditure.

5. *Regime PT4*: The last row of Table-I describes regime PT4. In this regime, when taking the tax evasion decision, consumer preferences are given by prospect theory, as in regimes PT1, PT2 and PT3. However, the government also uses these same prospect theory preferences to maximize the sum of the utilities of all the consumers. The details are given in section 4, below. As in regime EUT, we do not need to invoke either context dependence or paternalism. The results turn out not only to be empirically incorrect (as in regime EUT) but also economically absurd.

To summarize, in the context of the simple model of this paper, EUT cannot reconcile the observed tax rate, tax gap and the level of government expenditure. At the other extreme, modelling both the tax evasion decision and government behavior using PT gives absurd results. The best results are obtained by using PT to model tax evasion behavior but standard utility theory to model government behavior. The key is to recognize the context dependence of preferences.

3. The Taxpayer's Tax Evasion Problem: Expected Utility Theory

Consider the case of a consumer-taxpayer located at (x, s) who derives utility, $U(C, G)$, from private consumption, C , and the level of public provision, G . Given his income, $Y(x, s)$, the level of public provision, G , and the values of the parameters t , p , λ , and s , the consumer chooses the amount of income to declare, D . If he is not caught (with probability $1 - p$), then his disposable income, and hence private consumption, is Y_{NC} . However, if he is caught (with probability p), then his disposable income, hence private consumption, is Y_C . His expected utility is thus

$$EU = pU(Y_C, G) + (1 - p)U(Y_{NC}, G), \quad (3.1)$$

where Y_C and Y_{NC} are given by (2.13) and (2.12), respectively.²⁵

3.1. The Yitzhaki result

Eliminate D from (2.13) and (2.12) to get

$$\frac{1}{1 + \left(\frac{s}{t} + \lambda\right)} Y_C + \frac{\frac{s}{t} + \lambda}{1 + \left(\frac{s}{t} + \lambda\right)} Y_{NC} = (1 - t) Y. \quad (3.2)$$

For the special case of no stigma, $s = 0$, (3.2) reduces to

$$\frac{1}{1 + \lambda} Y_C + \frac{\lambda}{1 + \lambda} Y_{NC} = (1 - t) Y. \quad (3.3)$$

We may view the problem as choosing Y_C and Y_{NC} so as to maximize expected utility (3.1) subject to the budget constraint (3.3), given income $(1 - t)Y$ and prices $\frac{1}{1 + \lambda}$ and $\frac{\lambda}{1 + \lambda}$. Since prices do not depend on the tax rate, t , an increase in the tax rate has a pure income effect. Making the plausible assumption of constant or declining absolute risk aversion, we get that an increase in the tax rate reduces tax evasion.²⁶

In the more general case with stigma, (3.2), a change in the tax rate will have both income and substitution effects. However, simulations with plausible functional forms and parameter values indicate that in the presence of stigma, an increase in the tax rate causes a decline in evasion under EUT.²⁷

This result, obtained by Yitzhaki (1974), is rejected by the bulk of experimental, econometric and survey evidence.²⁸

²⁵Obviously, EU depends on D , G , x , s , t , λ , σ , s , and p . We have omitted reference to these to reduce the burden of notation.

²⁶For a formal proof see, for example, Dharami and al-Nowaihi (2007, Proposition 2). Along with constant or declining absolute risk aversion, we also need the assumptions $0 \leq D \leq Y$, $\left[\frac{\partial EU(D)}{\partial D}\right]_{D=0} > 0$, $\left[\frac{\partial EU(D)}{\partial D}\right]_{D=Y} < 0$, $\frac{\partial^2 EU(D)}{\partial D^2} < 0$, $p > 0$, $\lambda > 0$ and $t > 0$.

²⁷See below and Dharami and al-Nowaihi (2007).

²⁸See, for example, Friedland et al. (1978), Clotfelter (1983), Baldry (1987), Andreoni et al. (1998) and Pudney et al. (2000). However, a notable exception is Feinstein (1991).

3.2. Tax evasion under EUT

As described above in section 2.4, in the regime EUT when the taxpayer uses EUT, the preferences over private and public goods consumption are given by (2.15) and expected utility is given by (2.16). Note that expected utility (2.16) differs across consumer-taxpayers only in so far as they differ in income, $Y(x, s)$ and stigma, s . Differentiating (2.16) with respect to D , using (2.12) and (2.13), gives

$$\frac{\partial EU}{\partial D} = (1 - \mu) \left[\frac{1 - p}{Y_{NC}} \frac{\partial Y_{NC}}{\partial D} + \frac{p}{Y_C} \frac{\partial Y_C}{\partial D} \right], \quad (3.4)$$

$$\frac{\partial EU}{\partial D} = (1 - \mu) \left[(s + \lambda t) \frac{p}{Y_C} - t \frac{1 - p}{Y_{NC}} \right], \quad (3.5)$$

$$\frac{\partial^2 EU}{\partial D^2} = -(1 - \mu) \left[p \left(\frac{s + \lambda t}{Y_C} \right)^2 + (1 - p) \left(\frac{t}{Y_{NC}} \right)^2 \right] < 0. \quad (3.6)$$

If $t > 0$, then (3.6) holds everywhere. However, if $t = 0$, then $\frac{\partial^2 EU}{\partial D^2} = 0$ at $s = 0$. Hence, (3.6) holds almost everywhere. Since EU is continuous on the compact interval, $0 \leq D \leq Y(x, s)$, a maximum, $D(t, G, x, s)$, exists. This maximum is unique except at $t = s = 0$. The first order conditions for a maximum are

$$\left[\frac{\partial EU}{\partial D} \right]_{D=0} \leq 0, \quad \left[\frac{\partial EU}{\partial D} \right]_{0 < D < Y(x, s)} = 0, \quad \left[\frac{\partial EU}{\partial D} \right]_{D \geq Y(x, s)} \geq 0. \quad (3.7)$$

Let²⁹

$$s_1 = \min \left[1, \max \left(0, \frac{(1 - p)(t - t^2)}{(t - pt + p)} - \lambda t \right) \right], \quad (3.8)$$

$$s_2 = \max \left[s_1, \min \left(1, \frac{1 - p}{p} t - \lambda t \right) \right]. \quad (3.9)$$

From (2.12), (2.13), (3.5), (3.7), (3.8) and (3.9) we get

$$\begin{aligned} 0 \leq s \leq s_1 &\Rightarrow D = 0 \\ s_1 < s < s_2 &\Rightarrow D = \left[1 + \frac{p}{t} - p - \frac{(1-p)(1-t)}{s+\lambda t} \right] Y \\ s_2 \leq s \leq 1 &\Rightarrow D = Y \end{aligned} \quad (3.10)$$

Thus, taxpayers who would suffer low stigma if caught (the interval $0 \leq s \leq s_1$) hide all their income. Taxpayers who would suffer moderate stigma if caught (the interval $s_1 < s < s_2$) hide some but not all income. Taxpayers who would suffer high stigma if caught (the interval $s_2 \leq s \leq 1$) declare all their income. Also note that declared income, D , does not depend on the level of public provision, G . The latter result follows from the facts that utility (2.15) is additively separable in private and public consumption and the consumer-taxpayer takes public provision as given when deciding who much income to declare.

²⁹For $p > 0$ and $0 \leq t \leq 1$, it can be shown that $s_1 \leq \frac{1-p}{p}t - \lambda t$.

3.3. Tax revenue under EUT

From (3.10) we get:

$$\begin{aligned}
\int_{x=0}^1 \int_{s=0}^1 D dx ds &= \int_{x=0}^1 \int_{s=0}^{s_1} D ds dx + \int_{x=0}^1 \int_{s=s_1}^{s_2} D ds dx + \int_{x=0}^1 \int_{s=s_2}^1 D ds dx, \\
&= \int_{x=0}^1 \int_{s=s_1}^{s_2} \left(1 + \frac{p}{t} - p - \frac{(1-p)(1-t)}{s + \lambda t} \right) Y ds dx \\
&\quad + \int_{x=0}^1 \int_{s=s_2}^1 Y ds dx.
\end{aligned} \tag{3.11}$$

We now invoke the assumption that income distribution, $Y(x, s)$, is independent of stigma, s , so that $\int_{x=0}^1 Y(x, s) dx = \int_{x=0}^1 Y(x) dx = \bar{Y}$. Hence, (3.11) becomes

$$\begin{aligned}
\int_{x=0}^1 \int_{s=0}^1 D dx ds &= \bar{Y} \int_{s=s_1}^{s_2} \left(1 + \frac{p}{t} - p - \frac{(1-p)(1-t)}{s + \lambda t} \right) ds + (1 - s_2) \bar{Y}, \\
&= \left[1 - s_1 + (s_2 - s_1) \left(\frac{p}{t} - p \right) \right] \bar{Y} - (1 - p)(1 - t) \bar{Y} \int_{s=s_1}^{s_2} \frac{ds}{s + \lambda t}, \\
&= \left[1 - s_1 + (s_2 - s_1) \left(\frac{p}{t} - p \right) \right] \bar{Y} \\
&\quad - (1 - p)(1 - t) \bar{Y} [\ln(s_2 + \lambda t) - \ln(s_1 + \lambda t)].
\end{aligned} \tag{3.12}$$

Letting

$$\bar{D}(t) = \int_{x=0}^1 \int_{s=0}^1 D dx ds, \tag{3.13}$$

we get

$$\bar{D}(t) = \left[1 - s_1 + (s_2 - s_1) \left(\frac{p}{t} - p \right) \right] \bar{Y} - (1 - p)(1 - t) \bar{Y} [\ln(s_2 + \lambda t) - \ln(s_1 + \lambda t)]. \tag{3.14}$$

Substitute from (3.14) into (2.10), to get total tax revenue:

$$\bar{T}(t, G) = [\omega + p(1 + \lambda)(1 - \omega)] \bar{Y} t + (1 - \omega) [1 - p(1 + \lambda)] t \bar{D}(t). \tag{3.15}$$

Now, impose the government budget constraint, $\bar{T}(t, G(t)) = G(t)$, to get total public expenditure (which is also per capita public expenditure):

$$G(t) = [\omega + p(1 + \lambda)(1 - \omega)] \bar{Y} t + (1 - \omega) [1 - p(1 + \lambda)] t \bar{D}(t). \tag{3.16}$$

4. The Taxpayer's Tax Evasion Problem: Prospect Theory

The preferences of the taxpayer under prospect theory are more complicated. Those familiar with prospect theory should just skim the material in this section, skipping directly

to (4.9) for the value function under PT. For those unfamiliar with PT, we provide a self-contained treatment below.

The basic building blocks of prospect theory can be heuristically explained as follows.³⁰ Prospect theory distinguishes between two phases in decision making: an *editing phase*, followed by an *evaluation phase*. In the editing phase, a complex problem is first simplified to facilitate decision making. In the evaluation phase, the highest value prospect is chosen. During the editing phase outcomes are coded as *gains* or *losses* relative to a *reference point*. The reference point is usually, but not necessarily, the status quo.³¹

While there is no general theory of the editing phase, prospect theory has a very precise theory of the evaluation phase. Suppose that a consumer faces a lottery (or prospect) with several possible outcomes. First, each outcome in the prospect is assigned a number, using a *utility function*. This number is a positive real number if it has been coded as a gain relative to the reference point, and a negative number if it has been coded as a loss (the reference point having been arrived at in the editing phase). The utility function under prospect theory has the following properties: *continuity*, *monotonicity*, *reference dependence*, *declining sensitivity* and *loss aversion*. Continuity and monotonicity are as in EUT.

Unlike EUT where the carriers of utility are final levels of wealth (or incomes or consumption levels or commodities), under prospect theory the carriers of utility are gains and losses relative to the reference point. Declining sensitivity means that the utility function is concave in the domain of gains and convex in the domain of losses. Loss aversion is based on the idea that losses are more salient than gains. Given an amount of money, $y > 0$, and a utility function, $v(y)$, (to be specified precisely below) loss aversion implies that $v(y) > -v(-y)$.

Finally, the utilities of each outcome in a prospect are aggregated using *decision weights* into a *value function*. These decision weights are non-linear functions of the cumulative probabilities. Probabilities in the two domains (gains and losses) being cumulated separately. The *probability weighting function* used for the domain of gains need not be the same as that for losses. Decision weights are not probabilities and do not, necessarily, add up to one (unlike EUT). However, if all outcomes are either in the domain of gains or if all are in the domain of losses, then decision weights do add up to one and can be interpreted as probabilities. Agents facing uncertain situations overweight small probabilities but underweight large ones.³² In choosing among several prospects, an individual using

³⁰There is a substantial body of evidence in support of these building blocks of prospect theory, as well as a mounting number of successful applications in economics; see, for instance, the collection of papers in Kahneman and Tversky (2000) and in Camerer et al. (2004).

³¹Also in the editing phase it is decided which low probability events to ignore and which high probability events to treat as certain.

³²In Kahneman and Tversky (1979) decision weights are transformed probabilities. This proved unsatis-

PT chooses the one that gives rise to a higher number for the value function.

We now provide a more formal treatment of the building blocks of prospect theory.

4.1. Utility of an outcome under PT

Since the consumer takes public provision as given when deciding how much tax to evade, let us assume that public provision is ignored in the editing phase. Let the reference private consumption of the taxpayer be R . Then private consumption relative to the reference point is

$$X_i = Y_i - R, i = C, NC. \quad (4.1)$$

As in Kahneman and Tversky (1979) and Tversky and Kahneman (1992), the utility, $v(X_i)$, associated with an outcome X_i is given by

$$v(X_i) = \begin{cases} X_i^\beta & \text{if } X_i \geq 0, \\ -\theta(-X_i)^\beta & \text{if } X_i < 0, \end{cases} \quad (4.2)$$

where $\theta > 1$ is the parameter of loss aversion; it ensures that a loss is more salient than a gain of equal monetarily value. Based on experimental evidence, Tversky and Kahneman (1992) suggests that $\beta \simeq 0.88$ and $\theta \simeq 2.25$.³³

4.2. The reference point under PT

Although prospect theory does not provide sufficient guidance to determine the reference point in each possible situation, there is often a plausible candidate for a reference point. Indeed, specifying a suitable reference point is often essential for a successful application of prospect theory.

As in Dhami and al-Nowaihi (2007), we take the legal after-tax income (which is also the level of private consumption expenditure in the absence of tax evasion) as the reference point in this paper.³⁴ Hence,

$$R = (1 - t)Y. \quad (4.3)$$

factory. So, in Tversky and Kahneman (1992), decision weights are transformed cumulative probabilities. This produced cumulative prospect theory (CPT). However, since we have exactly one outcome in the domain of gains and exactly one outcome in the domain of losses, CPT reduces to PT. If non-linear transformation of cumulative probabilities is combined with EUT, we get *rank dependent expected utility* (RDEU). The attraction of RDEU is that it can be regarded as EUT applied to the transformed cumulative probability distribution. Hence, the whole machinery of analysis of risk developed for EUT can be transferred to RDEU. See Quiggin (1993) for details. While RDEU can explain some anomalies of EUT it cannot explain others. For the latter, we need PT. For application of RDEU to tax evasion, see Eide (2001).

³³See al-Nowaihi, Bradley and Dhami (2007) for an axiomatic derivation of the form (4.2).

³⁴Arguments for this particular reference point are given in Dhami and al-Nowaihi (2007). Alternative reference points are analyzed in Bernasconi and Zanardi (2004).

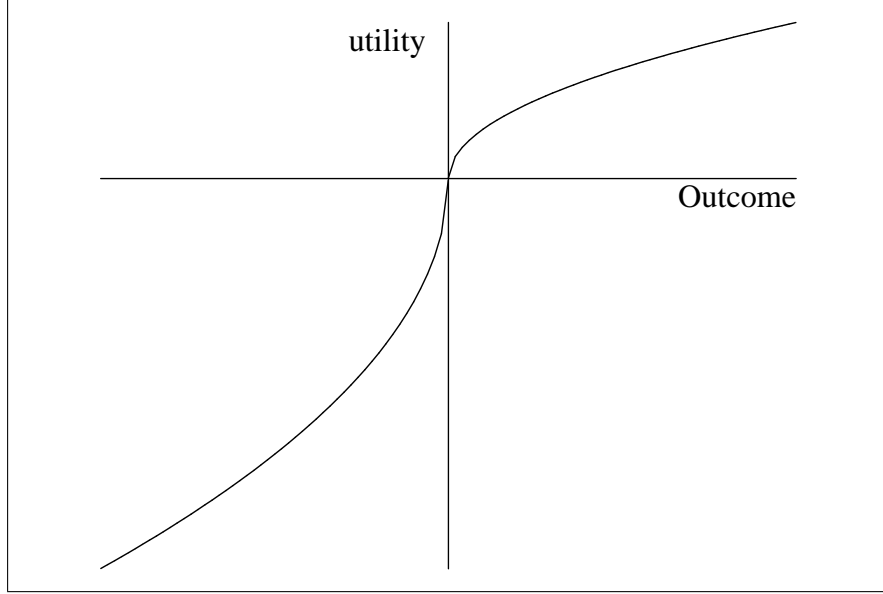


Figure 4.1: **Plot of prospect theory preferences for $\beta = 0.5$, $\theta = 2.5$**

4.3. The decision problem under PT

From (4.1), (4.3) $X_{NC} = Y_{NC} - (1 - t)Y$ and $X_C = Y_C - (1 - t)Y$. Then, using (2.12), (2.13) and recalling that $0 \leq D \leq Y$, we get

$$X_{NC} = t(Y - D) \geq 0, \quad (4.4)$$

$$X_C = -(s + \lambda t)(Y - D) \leq 0. \quad (4.5)$$

Hence, the taxpayer is in the *domain of losses* if caught but in the *domain of gains* if not caught. Let v be the taxpayer's value function and w^+ , w^- be her probability weighting function for the domains of gains and losses, respectively.³⁵ Then, according to *prospect theory* (PT), the taxpayer maximizes:

$$V = w^-(p)v(X_C) + w^+(1 - p)v(X_{NC}), \quad (4.6)$$

Comparing (4.6) with the analogous expression (3.1) for expected utility theory, we see the following differences. First, the carriers of utility in PT are gains and losses relative to the reference point rather than final levels. Second, one uses decision weights in PT to aggregate outcomes while one uses objective probabilities under expected utility theory.

³⁵By a *probability weighting function* we mean a strictly increasing function $w : [0, 1] \xrightarrow{onto} [0, 1]$. Note that a probability weighting function, w , has a unique inverse, $w^{-1} : [0, 1] \xrightarrow{onto} [0, 1]$ and that w^{-1} is strictly increasing. Furthermore, it follows that w and w^{-1} are continuous and must satisfy $w(0) = w^{-1}(0) = 0$ and $w(1) = w^{-1}(1) = 1$. For an example of a probability weighting function, see below.

Third, the level of public provision, G , is present in (3.1) but absent for (4.6). This is because, in EUT, a decision maker has a complete preference relation over all outcomes.

From (4.2), (4.4), (4.5) and (4.6) we get:

$$V = t^\beta (Y - D)^\beta w^+(1 - p) - \theta (s + \lambda t)^\beta (Y - D)^\beta w^-(p). \quad (4.7)$$

4.4. The probability weighting function

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 p the cumulative probability, Prelec (1998) derives the following weighting function (see Figure 4.2).³⁶ By the *Prelec function* we mean the probability weighting function $w : [0, 1] \xrightarrow{onto} [0, 1]$ given by³⁷:

$$w(p) = w^+(p) = w^-(p) = \exp [-(-\ln p)^\alpha], \quad (4.8)$$

where $0 < \alpha \leq 1$. The smaller α is, the more the overweighting of small probabilities

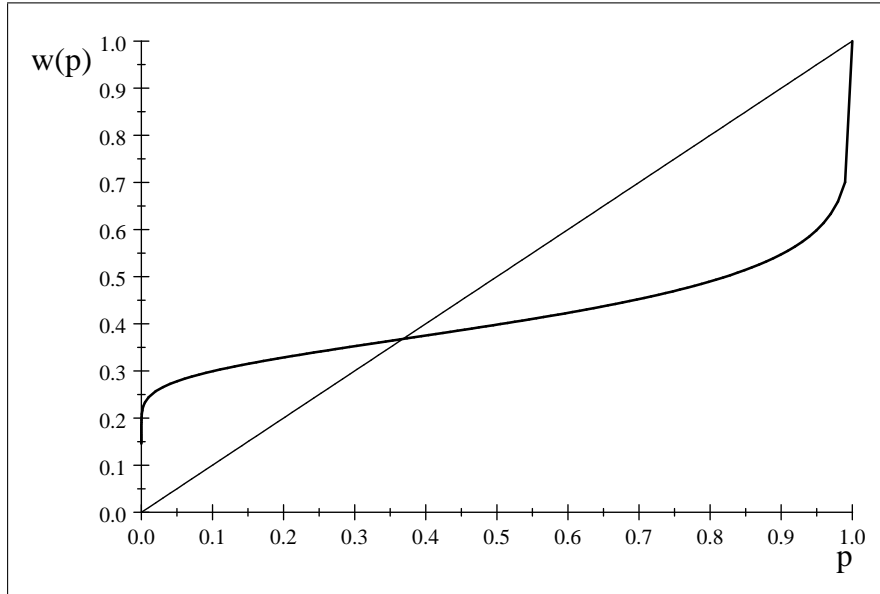


Figure 4.2: A graph of the Prelec weighting function for $\alpha = 0.225$

³⁶There are several advantages in using the Prelec weighting function, relative to the others suggested in the literature. First, it has an inverted S shape which is consistent with experimental evidence. Second, it is based on axiomatic foundations; see Prelec (1998), Luce (2001) and al-Nowaihi and Dhami (2006). Third, it has the same form for gains and losses.

³⁷To quote from Prelec (1998, last line of Appendix A): “Empirically ... one observes $w^+(p) = w^-(p)$.” Therefore, in our calibration exercises, we shall take $w^+(p) = w^-(p)$.

and $w(p) \rightarrow p$ as $\alpha \rightarrow 1$. Thus, the weights approach objective probabilities as $\alpha \rightarrow 1$.³⁸ Figure 4.2 plots the Prelec function for $\alpha = 0.225$.³⁹

4.5. Tax evasion under PT

Substituting (4.8) in (4.7) we get

$$V = t^\beta (Y - D)^\beta w^+(1 - p) - \theta (s + \lambda t)^\beta (Y - D)^\beta w^-(p). \quad (4.9)$$

Let

$$\psi = \left(\frac{w^+(1 - p)}{\theta w^-(p)} \right)^{\frac{1}{\beta}} - \lambda, \quad (4.10)$$

$$\begin{aligned} s_c &= 0 & \text{if } \psi t \leq 0 \\ s_c &= \psi t & \text{if } 0 < \psi t < 1 \\ s_c &= 1 & \text{if } \psi t \geq 1 \end{aligned} \quad (4.11)$$

then (4.9) can be written as:

$$V = \theta w^-(p) \left[(s_c + \lambda t)^\beta - (s + \lambda t)^\beta \right] (Y - D)^\beta. \quad (4.12)$$

From (4.12), we see that optimal values of D are

$$\begin{aligned} \text{Case-I: } D &= 0 & \text{if } s < s_c, \\ \text{Case-II: } D &\in [0, Y] & \text{if } s = s_c, \\ \text{Case-III: } D &= Y & \text{if } s > s_c. \end{aligned} \quad (4.13)$$

The solution to the tax evasion problem under PT, when the probability of detection is fixed, is a *bang-bang (or corner) solution*. We would argue that the *bang-bang solution* seems descriptive of several forms of tax evasion which take the form of hiding certain activities completely from the tax authorities while fully declaring other sources. For instance, an academic might not report income arising from an invited but paid lecture. A school teacher might not report tuition income for after-school lessons. A householder might pay cash to a builder for a minor extension of the house. Line item reporting of tax returns might further encourage this behavior. This, bang-bang, implication of reporting taxable income can also be drawn from the experimental results of Pudney

³⁸There is no reason to suppose that different uncertain situations should have the same α , hence, prospect theory does not impose an exact value on α . For instance, people might overweight the probability of dying in an air crash far higher relative to dying in a car accident.

³⁹Based on experimental data, Prelec (1998) estimates $\alpha \simeq 0.65$. However, Bernasconi (1998) argues that, because of ambiguity aversion, taxpayers ‘in the wild’ would exhibit more overweighting of low probabilities than in the laboratory. Bernasconi (1998) reports that, while actual probabilities of audits are in the range 0.01 to 0.03, an average of USA taxpayers’ assessments of the audit probability is 0.09. Taking a central value, this gives the value $\alpha = 0.225$ that we use.

et al. (2000). Slemrod and Yitzhaki (2002) find, based on TCMP data for 1988, that “the voluntary reporting percentage was 99.5% for wages and salaries, but only 41.4% for self-employment income”. Additional support comes from the behavior of non-profit organizations whose profits from activities unrelated to their primary tax exempt purpose are subject to federal and state tax. The reporting behavior of such organizations is also suggestive of the bang-bang solution; see Omer and Yetman (2002).⁴⁰

Assuming non-increasing absolute risk aversion, *EUT* predicts (Yitzhaki, 1974) that individuals evade *less* income as the tax rate *increases*. On the other hand, *PT* predicts the more factual result that tax evasion increases with an increase in the tax rate. This is formally stated in Proposition 1 below. The proof of this, and other results, can be found in Dhimi and al-Nowaihi (2007).

Proposition 1 : *Ceteris-paribus*, $\exists t = t_c \in [0, 1]$ such that the individual does not evade taxes if $t < t_c$ but evades taxes if $t > t_c$.

Consider the more general case of an endogenous probability of detection $p(D)$ such that $p(D)$ is continuously differentiable and $p'(D) \leq 0$. In this case the declared income can have an interior solution and varies continuously with the exogenous parameters. The comparative static results in this section, including, in particular, the explanation of the Yitzhaki puzzle, can also be demonstrated for the more general case.⁴¹ However, the general case is not very conducive for undertaking calibration exercises which is the main method we use for distinguishing among alternative theories.⁴²

⁴⁰Dhimi and al-Nowaihi (2007) introduce an endogenous probability of detection $p(D)$ such that $p(D)$ is continuously differentiable and $p'(D) \leq 0$ i.e. the taxpayer is more likely to be caught if s(he) evades more. This induces enough curvature in the model for interior solutions. However, this does not alter the comparative static results.

⁴¹For the general case, Dhimi and al-Nowaihi (2007, Proposition 4) prove the following: (a) At a regular interior optimum, tax evasion is strictly decreasing in the punishment rate, λ , the stigma rate, s , and the coefficient of loss aversion, θ . However, tax evasion is strictly increasing in the tax rate, t .

(b) At an optimum on the boundary ($D^* = 0$ or $D^* = Y$), tax evasion is non-increasing in the punishment rate, λ , the stigma rate, s , and the coefficient of loss aversion, θ . Tax evasion is non-decreasing in the tax rate, t .

⁴²Even in macroeconomics where calibration is most prevalent, despite the non-linearities it is usually the simplified log-linearized version of the model that is used in the calibration exercises.

4.6. Tax revenue under PT

From (4.13) we get:

$$\begin{aligned}
\overline{D}(t) &= \int_{x=0}^1 \int_{s=0}^1 D(t, G, x, s) dx ds \\
&= \int_{x=0}^1 \int_{s=0}^{s_c} D dx ds + \int_{x=0}^1 \int_{s=s_c}^1 D dx ds \\
&= \int_{x=0}^1 \int_{s=s_c}^1 Y dx ds.
\end{aligned} \tag{4.14}$$

Now, invoke our assumption that stigma is independent of pretax income, $Y(x, s) = Y(x)$. Then (2.11), (4.11) and (4.14), we get:

$$\overline{D}(t) = \begin{cases} \overline{Y} & \text{if } \psi t \leq 0 \\ (1 - \psi t) \overline{Y} & \text{if } 0 < \psi t < 1 \\ 0 & \text{if } \psi t \geq 1 \end{cases} \tag{4.15}$$

From (2.10), (4.15) and the government budget constraint, $G = \overline{T}(t, G)$, we get total tax revenue:

$$G(t) = \begin{cases} \overline{Y} t & \text{if } \psi t \leq 0 \\ \overline{Y} t [1 - (1 - \omega)(1 - p(1 + \lambda)) \psi t] & \text{if } 0 < \psi t < 1 \\ \overline{Y} t [\omega + p(1 + \lambda)(1 - \omega)] & \text{if } \psi t \geq 1 \end{cases} \tag{4.16}$$

5. Optimal tax under EUT (regime EUT)

We have, so far, considered the tax evasion decision of taxpayers when they respectively following EUT and PT and also the associated tax revenues under each form of preferences. We now analyze the optimal tax decision of a central planner under EUT.

The objective function of the government is explained in section 2.4 above, for the case of regime EUT. It chooses the tax rate so as to maximize the utility of a representative consumer, with average income, \overline{Y} , in the state where he does not evade⁴³:

$$U = \mu \ln [G(t)] + (1 - \mu) \ln [(1 - t) \overline{Y}], \tag{5.1}$$

given the government budget constraint (3.16), where $\overline{D}(t)$, s_1 and s_2 are given by (3.14), (3.8) and (3.9), respectively. We find the solution to this maximization problem by a simple search over values of $t \in [0, 1]$.

Calibration Values: We use the values $p = 0.015$ and $\lambda = 0.5$ that appear typical for the USA (Alm et al. (1992), Andreoni et al. (1998), Bernasconi (1998)). From Slemrod

⁴³Note that, like the family with 2.1 children, there may be no consumer in Ω with average income, who does not evade.

and Yitzhaki (2002) and Andreoni et al. (1998) we can infer the value $\omega = 0.4$. While the general view seems to be that stigma costs from evasion are low, for example see Brooks (2001), we are not aware of the exact magnitudes. Some evidence is available from stigma costs that arises from claiming welfare benefits. For Britain, Pudney et al. (2002) find that the total stigma costs (which they define as stigma, hassle, search costs, etc.) range from about 0.1 to 0.2. We do not know the appropriate value of μ , which determines society's preference for public compared to private consumption.

Calibration Strategy: Our strategy is as follows. For values of μ in the interval $(0, 1)$, we compute the optimal tax rate, the resulting level of tax revenue, $G(t)$, and the tax gap ratio, $\frac{t\bar{Y}-G(t)}{G(t)}$, where the denominator is total tax revenue and the numerator (the tax gap) is the difference between what is theoretically owed in taxes ($t\bar{Y}$) and what is actually collected in taxes ($G(t)$). For the USA, the tax gap ratio for 1992 was 0.222 and for 1998 it was 0.199 (Americans for Fair Taxation). For the USA, total government tax revenue, as a percentage of GDP was 32% for 2004 (Laurin, 2006). We use the normalization $\bar{Y} = 100$.

Calibration Results: The calibration results are tabulated in Table-II.

Table-II (Optimal tax, tax gap and government spending under EUT)

μ	0.1	0.2	0.3	0.4	0.41	0.5	0.6	0.65	0.7	0.8	0.9
t	0.10	0.22	0.37	0.52	0.54	0.66	0.73	0.77	0.80	0.87	0.94
s_1	0.73	0.62	0.43	0.21	0.19	0.002	0	0	0	0	0
s_2	1	1	1	1	1	1	1	1	1	1	1
$G(t)$	4.39	10.11	18.91	31.71	33.23	48.23	58.22	64.10	68.60	80.87	90.34
$\frac{t\bar{Y}-G(t)}{G(t)}$	1.28	1.18	0.94	0.64	0.61	0.37	0.25	0.20	0.17	0.09	0.04

From Table-II we see that as society's desire for public provision (measured by μ) increases, so does the optimal tax rate, t , and government tax revenue, $G(t)$. This result is exactly what is expected. However, tax evasion as measured by the tax gap ratio, $\frac{t\bar{Y}-G(t)}{G(t)}$, drops very dramatically with an increase in the tax rate, from the very high value of 1.28 to the very low value of 0.04. Thus tax evasion can be practically eliminated by a tax rate of almost 100%. This is the *Yitzhaki puzzle* under EUT described above.

From the fourth row we see that, at all tax rates, $s_2 = 1$ (see (3.9), (3.10)), i.e., all taxpayers who can evade, evade at least some tax. From the first column, we see that $s_1 = 0.73$ (see (3.8), (3.10)), i.e., of taxpayers who can evade, 73% evade all taxes. This steadily declines as the tax rate increases so that for tax rates in excess of 73%, all tax payers pay at least some tax.

From the first highlighted column, we see that at the tax rate $t = 0.52$, government tax revenue is similar to actual values ($G(0.52) = 31.71\%$, compared to the actual value of 32%). However, the predicted tax gap ratio is far too high ($\frac{0.52\bar{Y}-G(0.52)}{G(0.52)} = 0.64$, compared

to the actual value of about 0.2). From the second highlighted column, we see that the tax rate $t = 0.77$ gives a tax gap ratio close to that observed ($\frac{0.77\bar{Y}-G(0.77)}{G(0.77)} = 0.2$). However, the same tax rate gives a total tax revenue ($G(t) = 64.1\%$) that is about twice the observed value. To summarize, *our simple EUT optimal tax model with tax evasion can either get tax revenue right or the tax gap right but not both.*

We do not think that this result is merely the consequence of using a simple model, but appears to be a problem inherent in the EUT approach to tax evasion. We give some arguments in support of this:

1. (*Risk aversion*) The use of logarithmic utility entails a coefficient of relative risk aversion of 1, which is rather low. What results can we expect when a government assigns, in its social welfare function, a higher value for the coefficient of risk aversion? Bernasconi (1998) reports that coefficients of relative risk aversion lie in the range 1 to 2. However, Skinner and Slemrod (1985) report that a coefficient of relative risk aversion of 70 is needed to square the extent of tax evasion under EUT with the evidence.⁴⁴ Our simulations (see Section 7) support this, suggesting that increasing the coefficient of relative risk aversion to 2 would not make much difference.
2. (*More general preferences*) The simulations (see Section 7) also suggest that the additively separable form of the utility function for the representative taxpayer is not the problem, nor is the assumption of a representative consumer used to evaluate social welfare.
3. (*Stigma*) Increasing stigma cost would reduce tax evasion but, as discussed above, our assumed level of stigma appears already to be on the high side.
4. (*Other factors*) How about realistic features of the tax system that we have not included? For example, forms of taxes other than income taxes, fraudulent claims of benefits and costs of enforcing tax compliance. But in each of these cases we expect the paradox to reappear: at a tax rate that would generate the observed government tax revenue, EUT predicts too much evasion. Because of the Yitzhaki (1974) result, an increase in the tax rate that would reduce evasion to observed magnitudes would simultaneously increase tax revenue to well above observed values. Modelling labour supply has its own problems.⁴⁵

⁴⁴Mankiw and Zeldes (1991) calculated that a coefficient of relative risk aversion of 30 implies that the certainty equivalent of the prospect win 50,000 or 100,00 each with probability 0.5 is 51,209.

⁴⁵A large number of studies show a small *negative* aggregate labour supply elasticity of income (see, for example, Pencavel (1986) and Killingworth and Heckman (1986)). Clearly this would compound rather than solve the difficulties faced by EUT in explaining tax evasion. Many studies also show a small positive aggregate labour supply elasticity and some show a large positive elasticity. Many problems of applying EUT to labour supply are discussed in Bewley (1999).

6. Optimal tax under PT (regimes PT1 to PT4)

We now consider the four regimes, PT1 to PT4 in which the taxpayer uses PT in making the tax evasion decision. In all of them the tax evasion decision is modelled using prospect theory.

6.1. Regime PT1

Here the tax evasion decision is modelled using PT, as in section (4) above. In particular, tax revenues are now given by (4.16). However, the government's objective is still to maximize (5.1), the utility of a representative consumer-taxpayer with average income who does not evade. But now the relevant budget constraint is (4.16). It is routine, though tedious, to show that this optimizing problem has a unique maximum, and that the resulting optimal tax rate is in the interior of the interval $[0, 1]$. Hence, it can be found by solving the first order condition $U'(t) = 0$. Some simple algebra show that this is equivalent to solving $\mu(1-t) \frac{\partial G(t)}{\partial t} = (1-\mu)G(t)$. The solution is:

$$t = \mu, \text{ for } \omega = 1, \psi\mu \geq 1 \text{ or } \psi \leq 0. \quad (6.1)$$

If $\omega = 1$, then no one can evade. If $\psi \leq 0$ then all those who can evade decide not to. If $\psi\mu \geq 1$ then all those who can evade do so. In all these cases the optimal tax rate is $t = \mu$.

However, the more empirically relevant case is when $\omega < 1$ and $0 < \psi\mu < 1$. Here some taxpayers can evade but choose not to. In this case the first order condition gives the quadratic equation in t :

$$(1+\mu)(1-\omega)(1-p(1+\lambda))\psi t^2 - [1+2\mu(1-\omega)(1-p(1+\lambda))\psi]t + \mu = 0 \quad (6.2)$$

The optimum is the solution with the negative square root. It is given by:

$$t = \frac{1+2\mu(1-\omega)(1-p(1+\lambda))\psi - Z}{2(1+\mu)(1-\omega)(1-p(1+\lambda))\psi}, \quad \omega < 1, 0 < \psi\mu < 1,$$

where,

$$Z = \sqrt{[1+2\mu(1-\omega)(1-p(1+\lambda))\psi]^2 - 4\mu(1+\mu)(1-\omega)(1-p(1+\lambda))\psi}.$$

6.2. Regime PT2

The utility function (5.1) is additively separable over private and public provision. However, in actual practice, there are strong complementarities between the two. To take account of this, the utility function is now given by (2.17), as explained in section 2.4

under the regime PT2. In this model, we see that public provision has value only insofar as it facilitates private consumption. As with regime PT1, the government chooses the tax rate, t , so as to maximize the utility of a representative consumer with average income who does not evade, but now with the utility function:

$$U(t) = F\left((1-t)\bar{Y} - \frac{B}{G(t)}\right), F' > 0, F'' < 0, G(t) > 0, 0 < \frac{B}{G(t)} < (1-t)\bar{Y}, \quad (6.3)$$

where $G(t)$ is given by the budget constraint (4.16). The condition $0 < \frac{B}{G(t)} < (1-t)\bar{Y}$, together with (4.16), guarantees that any optimum must be an interior point. Hence it must satisfy the first order condition $U'(t) = 0$. Some simple algebra shows that this is equivalent to solving $B \frac{\partial G}{\partial t} = \bar{Y} G^2$. Using (4.16), the solution is

$$t = \begin{cases} \frac{\sqrt{B}}{\bar{Y}} & \text{for } \omega = 1 \text{ or } \psi \leq 0 \\ \frac{1}{\bar{Y}} \sqrt{\frac{B}{\omega + p(1+\lambda)(1-\omega)}} & \text{for } \psi t \geq 1 \end{cases} \quad (6.4)$$

If $\omega = 1$, then no one can evade. If $\psi \leq 0$ then all those who can evade decide not to. In both cases the optimal tax rate is $t = \frac{\sqrt{B}}{\bar{Y}}$. For the case $\psi t \geq 1$, all those who can evade, do evade. However, the most empirically relevant case is when $0 < \psi t < 1$. Here some taxpayers can evade but choose not to. In this case the first order condition gives the quartic equation in t :

$$\begin{aligned} & (1-\omega)^2 (1-p(1+\lambda))^2 \psi^2 \rho t^4 - 2(1-\omega)(1-p(1+\lambda)) \psi \rho t^3 + \rho t^2 \\ & + 2(1-\omega)(1-p(1+\lambda)) \psi t - 1 = 0; \quad 0 < \psi t < 1 \text{ and } \rho = \frac{(\bar{Y})^2}{B}. \end{aligned} \quad (6.5)$$

(6.5) has, of course, four roots. However, in simulations we always found the optimum to be the only root in the interval $(0, 1)$.

6.3. Regime PT3

So far, we have considered a representative consumer. An alternative welfare criterion is for the government to maximize the sum (or average) utility. For this we need to specify an income distribution. We choose the gamma distribution because of its tractability and because it gives quite a good representation of income distribution for middle incomes (which are the most relevant for determining aggregate tax revenue: the poor do not have much income and the rich are few in number) see, for example, Cowell (2000, p146). The Appendix gives all we need to know about the gamma distribution.

As explained in section 2.4 under the regime PT3, we generalize regime PT1 in two ways. Society's preference over private and public provision is given by the more general utility function (2.18). Instead of a representative consumer, the government now chooses

the tax rate, t , so as to maximize the sum (or average) utility, i.e., the social welfare function, $W(t)$, is now:

$$W(t) = \int_{x=0}^1 \int_{s=0}^1 \left[\mu \frac{[G(t)]^{1-\gamma}}{1-\gamma} + (1-\mu) \frac{[(1-t)Y]^{1-\gamma}}{1-\gamma} \right] dx ds, \quad (6.6)$$

where $G(t)$ is given by (4.16). Thus the government chooses the tax rate, t , so as to maximize society's total (or average) utility in the state where none evades tax. However, the government does allow for the effect on the tax revenue, $G(t)$, of tax evasion as described by PT. (6.6) simplifies to:

$$W(t) = \mu \frac{(G(t))^{1-\gamma}}{1-\gamma} + \frac{1-\mu}{1-\gamma} (1-t)^{1-\gamma} \int_{x=0}^1 \int_{s=0}^1 Y^{1-\gamma} dx ds. \quad (6.7)$$

The combination of the power form (2.18) for utility with the gamma distribution for income (10.1) leads to very tractable math. In particular, we get

$$\begin{aligned} \int_{x=0}^1 \int_{s=0}^1 Y^{1-\gamma} dx ds &= \frac{1}{b^a \Gamma(a)} \int_{Y=0}^{\infty} Y^{1-\gamma} Y^{a-1} e^{-\frac{Y}{b}} dY \\ &= \frac{b^{1-\gamma} \Gamma(1+a-\gamma)}{\Gamma(a)} \frac{1}{b^{1+a-\gamma} \Gamma(1+a-\gamma)} \int_{Y=0}^{\infty} Y^{(1+a-\gamma)-1} e^{-\frac{Y}{b}} dY \\ &= \frac{b^{1-\gamma} \Gamma(1+a-\gamma)}{\Gamma(a)} \end{aligned} \quad (6.8)$$

From (6.7) and (6.8), we get:

$$W(t) = \mu \frac{[G(t)]^{1-\gamma}}{1-\gamma} + \frac{1-\mu}{1-\gamma} (1-t)^{1-\gamma} \frac{b^{1-\gamma} \Gamma(1+a-\gamma)}{\Gamma(a)}. \quad (6.9)$$

Let

$$\varphi = \frac{1-\mu}{1-\gamma} (1-t)^{1-\gamma} \frac{b^{1-\gamma} \Gamma(1+a-\gamma)}{\Gamma(a)}.$$

From (4.16) and (6.9), we get:

$$W(t) = \begin{cases} \mu (1-\gamma)^{-1} (ab)^{1-\gamma} t^{1-\gamma} + \varphi & \text{if } \psi \leq 0 \\ \mu (1-\gamma)^{-1} (ab)^{1-\gamma} [t - (1-\omega)(1-p(1+\lambda))\psi t^2]^{1-\gamma} + \varphi & \text{if } 0 < \psi t < 1 \\ \mu (1-\gamma)^{-1} [\omega + p(1+\lambda)(1-\omega)]^{1-\gamma} (ab)^{1-\gamma} t^{1-\gamma} + \varphi & \text{if } \psi t \geq 1 \end{cases} \quad (6.10)$$

Maximizing $W(t)$ in the two cases $\psi \leq 0$ and $\psi t \geq 1$ is simple and leads to the optimal tax rates:

$$t = \begin{cases} \left[1 + \left(\frac{1-\mu}{\mu} \right)^{\frac{1}{\gamma}} \left(\frac{\Gamma(1-\gamma+a)}{a^{1-\gamma} \Gamma(a)} \right)^{\frac{1}{\gamma}} \right]^{-1} & \text{if } \psi \leq 0 \\ \frac{[\omega + p(1+\lambda)(1-\omega)]^{\frac{1-\gamma}{\gamma}}}{[\omega + p(1+\lambda)(1-\omega)]^{\frac{1-\gamma}{\gamma}} + \left(\frac{1-\mu}{\mu} \right)^{\frac{1}{\gamma}} \left(\frac{\Gamma(1-\gamma+a)}{a^{1-\gamma} \Gamma(a)} \right)^{\frac{1}{\gamma}}} & \text{if } \psi t \geq 1 \end{cases} \quad (6.11)$$

When $\psi \leq 0$, no taxpayer evades. When $\psi t \geq 1$, all those who can evade, do evade. However, the empirically more relevant case is when $0 < \psi t < 1$, where some who can evade do but others who can evade choose not to. For this case, the first order condition, $W'(t) = 0$ gives the following non-linear equation:

$$\begin{aligned} & [1 - 2(1 - \omega)(1 - p(1 + \lambda))\psi t]^{\frac{1}{\gamma}}(1 - t) \\ &= \left(\frac{1 - \mu}{\mu}\right)^{\frac{1}{\gamma}} \left(\frac{\Gamma(1 - \gamma + a)}{a^{1-\gamma}\Gamma(a)}\right)^{\frac{1}{\gamma}} [t - (1 - \omega)(1 - p(1 + \lambda))\psi t^2], \text{ if } 0 < \psi t < 1. \end{aligned} \quad (6.12)$$

Equation (6.12) can be easily solved numerically, given values for the parameters.

It is easy to check that, for $\gamma = 1$, (6.11) reduces to (6.1) and that (6.12) reduces to (6.2). Thus, the utilitarian regime, PT3, reduces to the representative consumer regime, PT1, for $\gamma = 1$.

6.4. Regime PT4

Finally, we consider a regime where tax evasion is described by PT and, also, the government chooses the tax rate, t , so as to maximize the sum (or average) of the PT utilities of the consumer-taxpayers. Thus, social welfare is now given by:

$$W(t) = \int_{x=0}^1 \int_{s=0}^1 V dx ds. \quad (6.13)$$

Substituting from (4.12) into (6.13), we get:

$$W(t) = \int_{x=0}^1 \int_{s=0}^1 \theta w^-(p) \left[(s_c + \lambda t)^\beta - (s + \lambda t)^\beta \right] [Y - D]^\beta dx ds. \quad (6.14)$$

From (4.13) and (6.14), we get

$$W(t) = \int_{x=0}^1 \int_{s=0}^{s_c} \theta w^-(p) \left[(s_c + \lambda t)^\beta - (s + \lambda t)^\beta \right] Y^\beta dx ds. \quad (6.15)$$

Invoking our assumption of independence of stigma and income we get, from (6.15),

$$\begin{aligned} W(t) &= \int_{s=0}^{s_c} \theta w^-(p) \left[(s_c + \lambda t)^\beta - (s + \lambda t)^\beta \right] ds \int_{x=0}^1 Y^\beta dx \\ &= \theta w^-(p) \left[s_c (s_c + \lambda t)^\beta - \frac{1}{(\beta + 1)} \left[(s_c + \lambda t)^{\beta+1} - (\lambda t)^{\beta+1} \right] \right] \int_{x=0}^1 Y^\beta dx. \end{aligned} \quad (6.16)$$

From (4.11) and (6.16), we get:

$$W(t) = \begin{cases} 0 & \text{if } \psi \leq 0 \\ \theta w^-(p) t^{\beta+1} \left[\psi (\psi + \lambda)^\beta - \frac{(\psi + \lambda)^{\beta+1} - \lambda^{\beta+1}}{1 + \beta} \right] \int_{x=0}^1 Y^\beta dx & \text{if } 0 < \psi t < 1 \\ \theta w^-(p) \left[(\lambda t)^\beta - \frac{(\lambda t)^{\beta+1} - (\lambda t)^{\beta+1}}{1 + \beta} \right] \int_{x=0}^1 Y^\beta dx & \text{if } \psi t \geq 1 \end{cases} \quad (6.17)$$

We shall consider only the case $0 < \psi t < 1$, since this is the empirically relevant one. Here, some taxpayers who can evade do, but others who can evade do not. It can be checked that the optimal tax depends on the sign of

$$\Delta = \left[\psi (\psi + \lambda)^\beta - \frac{1}{1 + \beta} \left[(\psi + \lambda)^{\beta+1} - \lambda^{\beta+1} \right] \right]. \quad (6.18)$$

If $\Delta < 0$, then welfare is strictly decreasing in t . Hence, the optimal tax rate is zero. This result is not surprising at all, since public provision does not figure in the utility functions, thus increasing the tax rate reduces private consumption without any gain. If $\Delta > 0$, then welfare is strictly increasing in the tax rate. Thus, the optimal tax rate is $t = 1$ and the ‘optimum’ is for the government to confiscate all the wealth of those who cannot evade or will not evade. This causes a huge loss in private consumption, but with no gain, since public provision is not in the utility functions.

The reason for this absurd result is that taxation shifts the reference point down. Those who cannot, or will not, evade are always at their reference point, so always have zero utility relative to that. Those who evade, and are not caught, have a huge increase in utility, measured *relative* to their reference point. But this huge relative increase in relative utility has no real welfare significance. Surprisingly, it is this, the most absurd of cases, that is picked up by our simulations (see next section). Hence, the message is that inferring preferences revealed in an inappropriate context can have disastrous policy effects.

7. Optimal Taxation: PT and EUT compared

We now compare the relative success of PT and EUT in explaining tax evasion. We use the same parameter values as before, when taxpayers had expected utility preferences: $p = 0.015$, $\lambda = 0.5$, and $\omega = 0.4$. We take the values $\beta = 0.88$ and $\theta = 2.25$ from Tversky and Kahneman (1992). We adopt the Prelec probability weighting function, $w^+ = w^- = e^{-(\ln p)^\alpha}$, with the value $\alpha = 0.225$ implied by the data reported by Bernasconi (1998). The results are tabulated below, where *na* stands for ‘not applicable’.

Table-III: Optimal taxes, spending and tax gaps in various regimes

Regime	μ	γ	B	t	s_1	s_2	s_c	$G(t)$	$\frac{t\bar{Y}-G(t)}{G(t)}$
EU	0.40	1	<i>na</i>	0.52	0.21	1	<i>na</i>	31.71	0.640
EU	0.65	1	<i>na</i>	0.77	0	1	<i>na</i>	64.10	0.200
PT1	0.44	1	<i>na</i>	0.387	<i>na</i>	<i>na</i>	0.282	32.28	0.198
PT2	<i>na</i>	<i>na</i>	1539	0.385	<i>na</i>	<i>na</i>	0.280	32.14	0.197
PT3	0.45	2	<i>na</i>	0.385	<i>na</i>	<i>na</i>	0.281	32.20	0.197
PT3	0.58	0.12	<i>na</i>	0.388	<i>na</i>	<i>na</i>	0.282	32.40	0.199
PT4	<i>na</i>	<i>na</i>	<i>na</i>	1	<i>na</i>	<i>na</i>	0.728	57.30	0.745

The first two rows are reproduced from Table-II, which gave results of calibrating regime EUT. Consumer-taxpayers are expected utility maximizers and use the same utility function to decide how much income to declare and also to express their preferences over private and public provision. The government uses this same utility function to determine the optimal tax rate. From the first row, we see that $\mu = 0.4$ gives about the correct level of total government tax revenues, $G(t)$, for the USA. However, the tax gap ratio, $\frac{t\bar{Y}-G(t)}{G(t)}$, is far too high, being about three times the correct value. On the other hand, from the second row, we see that $\mu = 0.65$ gives about the correct tax gap ratio for the USA. But then total government tax revenue is far too high, being about twice what it should be. The reason is that at observed audit probabilities and penalty and stigma rates, evasion is very attractive to an expected utility maximizer. Hence evasion is too high at observed levels of taxation. However, because of the Yitzhaki effect, *increasing* the tax rate *reduces* evasion. So, to get observed evasion rates, taxes have to be too high.

Rows 3 to 7 give the results when tax evasion is described by PT.

Row 3 gives the results for regime PT1. For PT1, consumer-taxpayers' evasion decisions are described by PT. However, their preferences over private and public consumption are given by the same utility function as in regime EUT.⁴⁶ The government chooses the tax rate so as to maximize the same welfare function as in regime EUT. However, the government takes into account the effect on tax revenues as predicted by PT.

Row 3 of the table shows that when consumer preferences over private and public consumption ($\mu = 0.44$) give a tax revenue ($G(t) = 32.28$) close to what is observed (32), we automatically get a tax gap ratio ($\frac{t\bar{Y}-G(t)}{G(t)} = 0.198$) close to what is observed (0.2). This suggests that PT provides the correct model for tax evasion.

The welfare function used in regime PT1 is additively separable over private and public consumption. By contrast, the welfare function used in regime PT2 exhibits strong complementarity between the two. From row 4, we see that when preferences between private

⁴⁶Alternatively, and as explained in the introduction, this could simply reflect paternalism on the part of the government. Our results do not depend on which interpretation is chosen. This comment applies to regimes PT2 and PT3 as well.

consumption and public provision ($B = 1539$) give a tax revenue ($G(t) = 32.14$) close to the observed value (32), they also gives a value for the output gap ratio ($\frac{t\bar{Y}-G(t)}{G(t)} = 0.197$) close to what is observed (0.2). Thus PT successfully explains tax evasion with two very different welfare functions.

For regime PT3, the criterion is to maximize aggregate (or average) utility, rather than that of a representative consumer of regimes PT1 and PT2. Also, regime PT3 generalizes PT1 in that it allows any (constant) coefficient of relative risk aversion, γ , while in PT1 this coefficient was $\gamma = 1$. In fact, for $\gamma = 1$, PT3 reduces to PT1. Since $\gamma = 1$ is at the low end of the range (between one to two) reported by Bernasconi (1998), in row 5 we choose $\gamma = 2$ to test the high end. The value $\mu = 0.45$ results in approximately the correct level of tax revenue ($G(t) = 32.2$) and gives a nearly correct output gap ratio (0.197). Row 6 reports the results for setting $\gamma = 0.12$, which is consistent with the value of $\beta = 0.88$, reported by Tversky and Kahneman (1992).⁴⁷ The value $\mu = 0.58$ gives a near correct value for tax revenue (32.4) and, again, a near correct value for the output gap ratio (0.199).

Using PT to model the tax evasion decision works for the following reasons. By taking the reference point to be the legal after-tax income, the taxpayer is in the domain of gains if not caught but in the domain of losses if caught.⁴⁸ This allows loss aversion, overweighting of the small probability of detection and underweighting the high probability of non-detection to considerably increase the deterrence effect of punishment, despite low values of p and λ (and the mild convexity of the value function for losses). This facilitates the derivation of the correct government budget constraint. By contrast, under EUT we get the wrong government budget constraint.

A fundamental principle of liberalism is that people are the best judges of their own welfare. However, the regimes PT1-PT3 use one set of preferences (PT) to model the tax evasion decision but another (standard utility theory) to describe preferences over private and public consumption. Furthermore, these regimes are successful in accounting for the observed facts on the tax gap and government expenditure.

This conclusion is further reinforced by the results of regime PT4. Regime PT4 uses PT to explain tax evasion (as in PT1-PT3), but *also* uses PT as a welfare criterion. Row 7 shows that PT4 gives wildly wrong values for all the variables. Worse, it gives the economically absurd result that when neither consumers nor government care about public provision, then the optimal tax rate is 100%. This absurd result comes about because PT4

⁴⁷The low value, $\gamma = 0.12$, should not be interpreted to mean that consumers are nearly risk neutral. For attitude to risk in PT depends, not only on the curvature of the value function, but also on the coefficient of loss aversion and the probability weighting function.

⁴⁸Proposition 3 of Dhimi and al-Nowaihi (2007) establishes that a consumer is in the domain of gains if caught but in the domain of losses if not caught if, and only if, the reference point is legal after tax income.

works by pushing down the reference point, so giving very high relative utility. But that does not correspond to any sensible measure of welfare. Thus, although PT gives the correct budget constraint in PT4, it also gives a completely wrong welfare criterion. The message is clear: inferring preferences from the wrong context can have disastrous policy effects.

8. Summary

A fundamental assumption of neoclassical economics is that decision makers have a complete transitive ordering over all possible outcomes. Research over the last 60 years has shown that this is not valid, not even as a rough approximation. With a simple general equilibrium model of optimal taxation in the presence of tax evasion, we were able to account well for the observed magnitudes. The key was recognizing that preferences are context dependent. In our successful regimes, we assumed that consumer-taxpayers behave according to prospect theory when they are considering how much income to declare. But that government, when maximizing social welfare on their behalf, use standard utility theory. However, when either expected utility on its own or prospect theory on its own is used to model both tax evasion and social welfare then the calibrated results are not in conformity with the data.

At observed audit probabilities and penalty and stigma rates, tax evasion is very attractive to an expected utility maximizer. Hence expected utility theory (EUT) predicts levels of evasion that are too high at observed levels of taxation. However, because of the Yitzhaki effect, *increasing* the tax rate *reduces* evasion. So, to get observed evasion rates, taxes have to be too high.

Under prospect theory (PT), taking the reference point to be the legal after-tax income ensures that the taxpayer is always in the domain of losses if caught and always in the domain of gains if not caught. Loss aversion, overweighting of low probabilities and the concavity of the value function for losses then ensure that punishment hurts more under PT than under EUT. Because of underweighting of high probabilities, the prospect of not being caught is less attractive under PT than under EUT. Hence, the deterrence effect of punishment is much stronger under PT. Thus, using PT to model tax evasion produces the observed levels of tax evasion.

Because of bounded rationality, consumer-taxpayers have to simplify considerably, concentrating on the salient features of the problem at hand. Hence, we cannot take their tax evasion behavior as revealing their preferences over private versus public consumption. Thus while PT gives the correct government budget constraint, it gives a completely wrong welfare criterion.

A fundamental principle of liberalism is that people are the best judges of their own

welfare. Neoclassical economics recognizes that applying this principle is problematic; governments have to take account of imperfections such as externalities, market failures and monopoly power. However, PT highlights additional issues. But, this is not necessarily a reason to restrict liberalism. On the contrary, by providing a descriptively more correct theory, PT can potentially enhance the applicability of liberal principles.

9. Acknowledgements

We are grateful for comments during invited seminar presentations at Cambridge, Warwick and Keele. In particular we would like to thank Joel Slemrod and Toke Aidt for comments and J. Barkley Rosser Jr. for the encouragement to write this paper.

10. Appendix: The Gamma Distribution

Most of the definitions, below, can also be found in Cowell (2000). The *gamma distribution* is given by

$$G(Y; a, b) = \frac{1}{b^a \Gamma(a)} \int_{y=0}^Y y^{a-1} e^{-\frac{y}{b}} dy, \quad a \geq 1, b > 0, \quad (10.1)$$

where $\Gamma(a)$ is the *gamma function*, given by:

$$\Gamma(a) = \int_{x=0}^{\infty} x^{a-1} e^{-x} dx, \quad a \geq 1. \quad (10.2)$$

If n is a positive integer, then

$$\Gamma(n) = (n-1)! \quad (10.3)$$

The mean and variance of the gamma distribution (10.1) are given by:

$$\bar{Y} = ab, \quad \text{var}Y = ab^2. \quad (10.4)$$

The *Gini index* for the gamma distribution is given by:

$$\begin{aligned} Gini(a, b) &= \frac{1}{\bar{Y}} \int_{Y=0}^{\infty} (2G(Y; a, b) - 1) Y dG(Y; a, b) \\ &= \frac{2b}{\bar{Y} (\Gamma(a))^2} \int_{Y=0}^{\infty} Y^a e^{-Y} \left(\int_{X=0}^Y X^{a-1} e^{-X} dX \right) dY - 1. \end{aligned} \quad (10.5)$$

The simplest gamma distribution that has the correct shape is:

$$G(Y; 2, 1) = \int_{y=0}^Y y e^{-y} dy, \quad (10.6)$$

for which

$$\begin{aligned}
Gini(2, 1) &= 0.375, \\
\text{mean} &= 2, \text{ median} = 1.6783, \\
\frac{\text{median}}{\text{mean}} &= \frac{1.6783}{2} = 0.83915.
\end{aligned} \tag{10.7}$$

Ryu and Slottje (2001, p303) report that estimates of Gini coefficients typically lie in the range 0.34 – 0.43, so $Gini(2, 1) = 0.375$ is not unreasonable.

For $G(Y; 2, 1)$, the *quantal functional*, $Q(p)$, the *cumulative income functional*, $C(p)$, and the *Lorenz curve*, $L(p)$, are given by, respectively,

$$\begin{aligned}
p &\in [0, 1], Y \in [0, \infty), p = G(Y; 2, 1) \Rightarrow Q(p) = Y, \\
C(p) &= \int_{y=0}^{Q(p)} y^2 e^{-y} dy = 2 - 2e^{-Q(p)} - 2Q(p)e^{-Q(p)} - (Q(p))^2 e^{-Q(p)}, p \in [0, 1], \\
L(p) &= \frac{C(p)}{\bar{Y}}, p \in [0, 1].
\end{aligned} \tag{10.8}$$

$Q(p)$ can be found by solving the equations $(1 - p)e^{x^2} = 1 + x^2$, $Q(p) = x^2$, where the square has been chosen to force the computer to the positive solution.

References

- [1] al-Nowaihi, A., Dhami, S., 2006. A simple derivation of Prelec's probability weighting function. *Journal of Mathematical Psychology* 50 (6), 521-524.
- [2] al-Nowaihi, A., Bradley, I., Dhami, S., 2007. Preference homogeneity and the power function in prospect theory. *Forthcoming in Economics Letters*.
- [3] Allingham, M.G., Sandmo, A., 1972. Income tax evasion: a theoretical analysis. *Journal of Public Economics* 1, 323-338.
- [4] Alm, J., McClelland, G.H., Schulze, W.D., 1992. Why do people pay taxes? *Journal of Public Economics* 48, 21-38.
- [5] Americans for Fair Taxation. The fair tax, tax evasion, and the underground economy. <http://www.fairtax.org/impact/evade24.html>.
- [6] Andreoni, J., Erard, B., Feinstein, J., 1998. Tax Compliance. *Journal of Economic Literature* 36, 818-860.
- [7] Atkinson, A.B., Stiglitz, J.E., 1972. The structure of indirect taxation and economic efficiency. *Journal of Public Economics* 1, 97-119.

- [8] Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: direct versus indirect taxation. *Journal of Public Economics* 6, 55-75.
- [9] Baldry, J.C., 1987. Income tax evasion and the tax schedule: some experimental results. *Public Finance* 42, 355-384.
- [10] Bartle, R.G., 1966. *The Elements of Integration*. John Wiley and Sons, Inc., New York. London. Sydney.
- [11] Bernasconi, M., 1998. Tax evasion and orders of risk aversion. *Journal of Public Economics* 67, 123-34.
- [12] Bernasconi, M., Zanardi, A., 2004. Tax evasion, tax rates and reference dependence. *FinanzArchiv* 60, 422-445.
- [13] Besley, T., Coate, S., 1992. Understanding welfare stigma: taxpayer resentment and statistical discrimination. *Journal of Public Economics* 48, 165-183.
- [14] Bewley, T.F., 1999. *Why wages don't fall during a recession*. Harvard University Press. Cambridge, Massachusetts, and London, England.
- [15] Boadway, R., Marchand, M., Pestieau, P., 1994 Towards a theory of direct-indirect tax mix. *Journal of Public Economics* 55, 71-88.
- [16] Boadway, R.W., Bruce, N., 1984. *Welfare Economics*. Basil Blackwell, England.
- [17] Brooks, N., 2001. Key issues in income tax: challenges of tax administration and compliance. In: Paper presented at the 2001 Tax conference of the Asian Development Bank
- [18] Camerer, C., Issacharoff, S., Loewenstein, G., O'Donoghue, T., Rabin, M. 2003. Regulation for conservatives: behavioral economics and the case for "asymmetric paternalism". *University of Pennsylvania Law Review* 151, 2111-1254.
- [19] Camerer, C.F., Loewenstein, G., Rabin, M., 2004. *Advances in Behavioral Economics*. Princeton University Press, Princeton.
- [20] Chandar, P., Wilde, L.L., 1998 A general characterization of optimal income tax enforcement. *Review of Economic Studies* 65, 165-83.
- [21] Clotfelter, C.T., 1983. Tax evasion and tax rates: an analysis of individual returns. *Review of Economics and Statistics* 65, 363-373.

- [22] Cowell, F.A., 2000. Measurement of inequality. In: Atkinson, A.B., Bourguignon, F. (Eds.), *Handbook of Income Distribution* vol.1. Elsevier Science, Amsterdam, pp. 87-166, Chapter 2.
- [23] Cremer, H., Gahvari, F., 1993. Tax evasion and optimal commodity taxation. *Journal of Public Economics* 50, 261-75.
- [24] Cremer, H., Gahvari, F., 1996. Tax evasion and optimal general income tax. *Journal of Public Economics* 60, 235-49.
- [25] D'Aspremont, C., Gevers, L., 1977. Equity and the informational basis of social choice. *Review of Economic Studies* 46, 199-210.
- [26] Dhami, S., al-Nowaihi, A., 2006. A simple model of optimal tax systems: taxation, measurement and uncertainty. *Manchester School* 74, 645-669.
- [27] Dhami, S., al-Nowaihi, A., 2007. Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior and Organization*.
- [28] Eide, E., 2001. Rank dependent expected utility models of tax evasion. Mimeo., University of Oslo.
- [29] Elffers, H., Hessing, D. J., 1997. Influencing the prospects of tax evasion. *Journal of Economic Psychology* 18, 289-304.
- [30] Feinstein, J.S., 1991. An econometric analysis of income tax evasion and its detection. *RAND Journal of Economics* 22, 14-35.
- [31] Friedland, N., Maital, S., Rutenberg, A., 1978. A simulation study of income tax evasion. *Journal of Public Economics* 10, 107-116.
- [32] Gordon, J.P.F., 1989. Individual morality and reputation costs as deterrents to tax evasion. *European Economic Review* 33, 797-805.
- [33] Harsanyi, J.C., 1953. Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy* 61, 434-435.
- [34] Harsanyi, J.C., 1955. Cardinal welfare, individualistic ethics and interpersonal comparisons of utility. *Journal of Political Economy* 73, 309-321.
- [35] Harsanyi, J.C., 1975. Can the maximin principle serve as a basis for morality? A critique of John Rawls' Theory. *The American Political Science Review* 69, 594-606.

- [36] Jolls, C., Sunstein, C. R., and Thaler, R., 1998. A Behavioral Approach to Law and Economics. *Stanford Law Review* 50, 1471-1541.
- [37] Kahneman, D., Tversky, A., 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47, 263-91.
- [38] Kahneman, D., Tversky, A. (Eds.), 2000. *Choices, Values and Frames*. Cambridge University Press, Cambridge.
- [39] Killingworth, M. R., Heckman, J.J., 1986. Female labor supply: a survey. In: Ashenfelter, O., Layard, R. (Eds.), *Handbook of Labor Economics*, vol.1. Elsevier Science, Amsterdam, pp. 103-204, Chapter 2.
- [40] Laurin, A., 2006. International tax burdens: Single individuals with or without children. PRB 05-107E. Library of Parliament - Parliamentary Information and Research Service.
- [41] Loeb, P.A., Wolff, M. (Eds.), 2000. *Nonstandard Analysis for the Working Mathematician*. Kluwer Academic Publishers.
- [42] Luce, R.D., 2001. Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology* 45, 167-179.
- [43] Mankiw, N.G., Zeldes, S., 1991. The consumption of stockholders and non stockholders. *Journal of Financial Economics* 29, 97-112.
- [44] Marhuenda, F., Ortuno-Ortin, I., 1997. Tax enforcement problems. *Scandinavian Journal of Economics* 99, 61-72.
- [45] Mas-Colell, A., Whinston, M.D., Green, J.R., 1995. *Microeconomic Theory*. Oxford University Press, New York, Oxford.
- [46] Maskin, E., 1978. A theorem on utilitarianism. *Review of Economic Studies* 45, 93-96.
- [47] Omer, T.C., Yetman, R.J., 2002. Near zero taxable income reporting by nonprofit organizations. In: *Fall Journal of the American Taxation Association*, vol. 25, no.2.
- [48] Pencavel, J., 1986. Labor supply of men: a survey. In: Ashenfelter, O., Layard, R. (Eds.), *Handbook of Labor Economics*, vol.1. Elsevier Science, Amsterdam, pp. 3-102, Chapter 1.
- [49] Prelec, D., 1998. The probability weighting function. *Econometrica* 60, 497-528.

- [50] Pudney, S.E., Pyle, D., Saruc, T., 2000. Income tax evasion: an experimental approach. In MacDonald, A., Pyle, D. (Eds.), *Illicit activity: The Economics of Crime and Tax Fraud*. Ashgate, Aldershot, pp267-289.
- [51] Pudney, S.E., Hernandez, M., Hancock, R., 2002. The welfare cost of means-testing: pensioner participation in income support. Discussion Papers in Economics 03/2, Department of Economics, University of Leicester.
- [52] Quattrone, G.A., Tversky, A., 1988. Contrasting rational and psychological analyses of political choice. *American Political Science Review* 82, 719-736.
- [53] Quiggin, J., 1993. Generalized expected utility theory: the rank-dependent expected utility model. Amsterdam: Kluwer-Nijhoff.
- [54] Richter, W.F., Boadway, R., 2006. Trading off tax distortion and tax evasion. Forthcoming in *Journal of Public Economic Theory*.
- [55] Rawls, J., 1971. *A Theory of Justice*. Harvard University Press, Cambridge, Mass.; and Clarendon Press, Oxford.
- [56] Ryu, H.K., Slottje, D.J., 2001. Parametric approximations of the Lorenz curve. In: Silber, J. (Ed.), *Handbook of Income Inequality Measurement*. Kluwer Academic Publishers, Boston/Dordrecht/London, pp. 291-314, Chapter 10.
- [57] Rubinstein, A., 2005. Comments on behavioral economics. In: Blundell, R., Newey, W.K., Persson, T. (Eds.), *Advances in Economic Theory vol.2*, Cambridge University Press, pp. 246-254.
- [58] Skinner, J., Slemrod, J., 1985. An economic perspective on tax evasion. *National Tax Journal* 38, 345-353.
- [59] Slemrod, J., 2007. Cheating ourselves: the economics of tax evasion. *Journal of Economic Perspectives* 21, 25-48.
- [60] Slemrod, J., Yitzhaki, S., 2002. Tax avoidance, evasion and administration. In: Auerbach, A.J., Feldstein, M. (Eds.), *Handbook of Public Economics vol.3*. Elsevier Science, Amsterdam, pp. 1423-1470, Chapter 22.
- [61] Tversky, A., Kahneman, D., 1986. Rational choice and the framing of decisions. *The Journal of Business* 59, S252-S278.
- [62] Tversky, A., Kahneman, D., 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297-323.

- [63] Yaniv, G., 1999. Tax compliance and advance tax payments: a prospect theory analysis. *National Tax Journal* 52, 753-764.
- [64] Yitzhaki, S., 1974. A note on income tax evasion: a theoretical analysis. *Journal of Public Economics* 3, 201-202.