

Prospect theory and tax evasion: a reconsideration of the Yitzhaki puzzle

Amedeo Piolatto¹  · Matthew D. Rablen² 

Published online: 14 November 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract The standard expected utility (EUT) model of tax evasion predicts that evasion is decreasing in the marginal tax rate (the Yitzhaki puzzle). Recent literature shows cases in which incorporating prospect theory (PT) does and does not overturn the Puzzle. In a general environment that nests both PT and EUT preferences, we provide a detailed study of how the elements of PT affect the Puzzle. PT does not always reverse the Puzzle, hence we give and interpret conditions for when it does and does not. When allowing for stigma and/or variable audit probability, PT reverses the Puzzle in the same way and with the same limitations as does EUT, if equally augmented.

Keywords Prospect theory · Tax evasion · Yitzhaki puzzle · Stigma · Diminishing sensitivity · Reference dependence · Variable audit probability · Endogenous reference level

JEL Classification H26 · D81 · K42

✉ Matthew D. Rablen
m.rablen@sheffield.ac.uk

Amedeo Piolatto
piolatto@ub.edu
<https://sites.google.com/site/piolatto>

¹ Barcelona Economics Institute (IEB), University of Barcelona, c/ J. M. Keynes, 1-11, 08034 Barcelona, Spain

² Department of Economics, University of Sheffield, 9 Mappin Street, Sheffield S1 4DT, UK

1 Introduction

The expected utility theory (EUT) model of tax evasion predicts a negative relationship between tax rates and evasion whenever fines are imposed on the evaded tax and taxpayers exhibit decreasing absolute risk aversion (Yitzhaki 1974). Although empirical evidence on this question is mixed, and depends to an extent on the econometric methodology used (Bernasconi et al. 2014), a substantial body of empirical and experimental evidence finds a positive relationship between evasion and the tax rate (see, e.g., Ali et al. 2001; Alm et al. 1995; Clotfelter 1983; Crane and Nourzad 1986; Friedland et al. 1978; Pommerehne and Weck-Hannemann 1996; Poterba 1987).¹ Owing to the weight of contradicting empirical evidence, and its counter-intuitive nature, the negative relationship between tax rates and evasion predicted by the EUT model has sometimes been termed the “Yitzhaki paradox” or “Yitzhaki puzzle”.

Prospect Theory (PT) has become a centrepiece of behavioural economics, for it is able to resolve many puzzles associated with EUT and it provides a better fit to much empirical data (Bruhin et al. 2010).² It remains disputed, however, whether the application of PT to tax evasion resolves the Yitzhaki puzzle: Bernasconi and Zanardi (2004), Dhami and al-Nowaihi (2007) and Yaniv (1999), among others, each illustrate a specification of PT that resolves the Yitzhaki puzzle, but Hashimzade et al. (2013, p. 16) show alternative specifications that do not. Unpicking these divergent results is far from straightforward, however, as the elements of PT are specified differently across studies and as some studies invoke auxiliary assumptions (in addition to those of PT). Hashimzade et al. (2013) do consider some general specifications of PT, but do not directly compare to other preference models, while Dhami and al-Nowaihi (2007) undertake a comparison of PT and EUT, but not in a unified framework that nests EUT and PT, and for only one specification of PT. In this paper, therefore, we seek to evaluate, in a general sense, the marginal contribution of the elements of PT—individually and collectively—towards resolving the Yitzhaki puzzle. By disentangling the separate driving forces, we are able to reconcile seemingly contradictory results in the literature and to clarify which of the elements of PT, if any, contribute to solving the Puzzle. As these elements are now widely applied in the broader literature on behavioural decision making, by isolating the different components, our results can be readily extended to many further behavioural models besides PT.

We perform our evaluation in a general environment—which we shall term the Taxpayer behavioural model (TBM)—in which it is possible to vary (i) the specification of reference income; (ii) the elements of PT that are assumed to hold—to separate out the distinct effects of reference dependence, diminishing sensitivity, loss aversion, and probability weighting; (iii) the auxiliary assumptions assumed to hold vis-à-vis those of the standard portfolio model of tax compliance. The TBM is sufficiently general to

¹ See Feinstein (1991), however, for evidence consistent with a negative relationship.

² PT was initially proposed by Kahneman and Tversky (1979), and subsequently extended to “cumulative” PT by Tversky and Kahneman (1992). In this study, we use cumulative PT, but our main qualitative conclusions apply also to the original version of PT. See, e.g., Barberis (2013) and Camerer (2000) for reviews of applications of PT other than to tax evasion.

encompass much of the existing literature, but sufficiently specific to yield conditions with clear economic and psychological interpretation.³

Our first contribution is to show that several seemingly disparate approaches to the specification of reference income in the existing literature are variants of a simple, yet general, formulation. Within this general formulation, our two main results are as follows: first, matching Yitzhaki's original demonstration of the Puzzle under EUT, we give apparently plausible conditions under which the Puzzle still holds under PT (and stripped-down variants). Second, we find that, although some specifications of PT do reverse the Puzzle, such reversals often rely on the psychologically questionable implication that a tax rise makes taxpayers feel subjectively richer (relative to reference income) in the not-caught state, and in expectation. In particular, for PT to resolve the Puzzle, this condition *must* hold when preferences are homogeneous, a common assumption in applications of PT. Thus, while our results do not necessarily endorse the descriptive validity of EUT, we find nonetheless that a set of specifications of PT—which includes many specifications proposed in the literature—is either psychologically questionable, or share similar descriptive deficiencies in respect of the Puzzle.

We examine the implications for both EUT and PT (among other variants) of allowing for two auxiliary assumptions: social stigma costs and a variable audit probability. Allowing for sufficient social stigma always resolves the Puzzle under EUT, but not always under PT. In contrast, allowing for a variable audit probability does not clearly improve the ability of either EUT or PT to resolve the Puzzle. In general, the conditions under which these auxiliary assumptions improve the predictions of PT with respect to the Puzzle are the same as those which also improve the predictions of EUT.

By allowing for stripped-down variants of PT, we observe the marginal contribution of each of its elements. Under PT preferences, reference dependence is necessary to overturn the Yitzhaki puzzle, but the remaining PT elements—diminishing sensitivity, loss aversion and probability weighting—are neither necessary nor sufficient. These findings seem consistent with the nascent literature on the relative economic importance of the PT elements—reference dependence being the most widely accepted (see, e.g., Barberis 2013; Santos-Pinto et al. 2015).

The results of this study contribute to the literature on the use of non-expected utility preferences (and PT in particular) to explain tax evasion, and to the wider literature on the descriptive usefulness of non-expected utility preferences (Kim 2005; Harrison and Rutström 2009; Bruhin et al. 2010; Isoni 2011; Rees-Jones 2014; Masatlioglu and Raymond 2016). We do not claim that EUT is descriptively superior or inferior to PT over the full gamut of empirical regularities on tax-related behaviour, and other evidence relating to behaviour in risky settings more generally. Our results do, though,

³ Among others, the TBM allows us to treat as special cases the models developed by Bernasconi and Zanardi (2004), Dhimi and al-Nowaihi (2007) and two cases described in Hashimzade et al. (2013). Our model can be used to reconcile the apparently contradictory results therein. Although we have strived for widespread applicability, the TBM does not exhaustively encompass the literature. In particular, Rablen (2010) introduces PT into a version of the tax evasion model that allows for taxes to fund the provision of a public good (which we do not consider here). In addition, Bernasconi et al. (2014) allow for reference income to adapt over time to changes in the tax rate. These authors show that their model predicts an upward drift in tax evasion (after an initial fall), following an increase in the tax rate.

lead us to question the claim that PT is inherently better able to reconcile the Yitzhaki puzzle than is EUT.

The rest of the paper is organised as follows. Section 2 introduces a general model of the tax evasion decision that nests both PT and EUT preferences. In Sect. 3, we analyse the model for a fixed audit probability, and then for a variable audit probability. Section 4 concludes with a discussion of our findings and some wider thoughts on the choice of reference income and of preferences in applications of PT to tax evasion. All proofs are given in the Appendix.

2 Taxpayer behavioural model

We now present a general model of the taxpayer tax evasion decision—the Taxpayer Behavioural Model—which nests PT, EUT and intermediate variants that may be considered as stripped-down versions of PT. Consider a taxpayer with an exogenous taxable income $Y > 0$ (which is known to the taxpayer but not to the tax authority).⁴ The government levies a proportional income tax at marginal rate $t \in (0, 1)$ on declared income $X \in (0, Y)$. The probability of audit is given by $p = p(X) \in (0, 1)$ with $p' \leq 0$, though taxpayers may behave as if they transform this objective probability into a decision weight $w(p)$, where $w(0) = 0$, $w(1) = 1$ and $w' > 0$.⁵ In the special case $p' = 0$ the probability of audit is independent, or “exogenous”, of declared income. Audited taxpayers face a fine at rate $f \in (1, t^{-1})$ on all undeclared tax, where the upper bound ensures that the amount paid (tax plus fines) never exceeds a taxpayer’s total income (hence limited liability is never violated, regardless of earned and declared income). Following [Dhami and al-Nowaihi \(2007\)](#), taxpayers may additionally incur social stigma equivalent to a monetary cost $s[Y - X]$, where $s \geq 0$. Accordingly, $Y^n \equiv Y - tX$ is the taxpayer’s income when not caught, and $Y^c \equiv Y^n - [tf + s][Y - X]$ is the taxpayer’s income when caught (audited).

The carrier of utility, v , is constructed to be sufficiently general to nest both EUT and PT preferences. Specifically, we first define a function $u : \mathbb{R} \mapsto \mathbb{R}$ satisfying

- A0. $u(x)$ is continuous and twice differentiable for all $x \in \mathbb{R}$, except possibly at $x = 0$.
- A1. $u(0) = 0$;
- A2. $u' > 0$;

⁴ Following the study of [Kleven et al. \(2011\)](#)—who show that the descriptive validity of the EUT model is improved when allowing for plausible levels of third-party reporting—it is now common to interpret Y not as a taxpayer’s full income, but as that part of income not subject to third-party reporting. As this interpretation is not adopted, however, in the literature we address, we do not formally adopt this interpretation.

⁵ Hence, the objective probability distribution is $\{p, 1 - p\}$ and the transformed probability distribution is $\{w(p), 1 - w(p)\}$. PT allows for different weighting functions to apply to outcomes that fall above or below the reference level. As pointed out by [Dhami and al-Nowaihi \(2007\)](#) and [Prelec \(1998\)](#), however, empirically the same weighting function is found to apply above and below the reference level, so we assume there to be a single weighting function w . The assumption $p' \leq 0$ is consistent with the literature on optimal auditing (e.g., [Reinganum and Wilde 1986](#); [Piolatto and Trotin 2016](#)).

- A3. $u'' < 0$;
- A4. $\partial A^u(x) / \partial x \leq 0$, where $A^u(x) \equiv -u''(x) / u'(x)$.⁶

The carrier of utility, v , is then defined as

$$v(x) = \begin{cases} u(x) & \text{if } x \geq 0; \\ [-1]^{DS} \lambda u([-1]^{DS} x) & \text{otherwise;} \end{cases} \tag{1}$$

where $\lambda \geq 1$, and the condition DS is true ($DS = 1$) when diminishing sensitivity is assumed to hold, and false otherwise ($DS = 0$). Note, first, that setting $DS = 0$ and $\lambda = 1$ we have $v(x) = u(x)$ as under EUT. In this case, A2 and A3 imply that v is increasing and concave. A4 implies that absolute risk aversion with respect to v is decreasing, although we note that for $w(p) \neq p$ risk preferences (in the sense of preferences for mean-preserving spreads) are determined jointly with respect to v and to w (Schmidt and Zank 2008).⁷ Note, second, that under diminishing sensitivity utility becomes $v(x) = -\lambda u(-x)$ for $x < 0$. In this case, A2 still implies that v is monotonic, but A3 implies that v is convex for $x < 0$. Following Köbberling and Wakker (2005), loss aversion with respect to v requires $\lim_{x \uparrow 0} \partial v(x) / \partial x > \lim_{x \downarrow 0} \partial v(x) / \partial x$, which holds if and only if $\lambda > 1$. Assumption A4 implies with respect to v that

$$A(x) \equiv A^v(x) \equiv -\frac{v''(x)}{v'(x)} \begin{cases} < 0 & \text{if } x < 0, \quad DS = 1; \\ > 0 & \text{otherwise.} \end{cases}$$

Taxpayers are assumed to judge outcomes relative to a reference level of income R —the reference dependence element of PT. Thus, we write $\Delta Y^i \equiv Y^i - R, i = c, n$. We employ a specification of reference income that we shall show is sufficiently general to nest a wide range of those proposed in the existing literature. The underlying observation informing our specification is that existing approaches can be understood as weighted averages of wealth in the caught and not-caught states. Specifically, we write

$$R = \alpha(\cdot) Y^c + [1 - \alpha(\cdot)] Y^n, \tag{2}$$

where the function $\alpha(\cdot)$ satisfies

$$A5. \alpha(\cdot) = \phi_1 + \frac{\phi_2}{Y - X},$$

and $\phi_1(\cdot)$ and $\phi_2(\cdot)$ are real-valued functions satisfying $\phi_{1X} = \phi_{2X} = 0$ (where $\phi_{iX} \equiv \partial \phi_i(\cdot) / \partial X$). For future reference, we denote the elasticity of ϕ_i with respect to t as $\varepsilon_{\phi_i,t}$. The specifications of reference income in the existing literature that are special cases of A5 are listed in Table 1. $R = 0$, as under EUT, is the special case of A5 obtained by setting $\phi_1 = t / [ft + s]$ and $\phi_2 = Y [1 - t] / [ft + s]$.

⁶ An example of a function satisfying A0-A4 is $u(x) = x - e^{-bx}$ with $b > 0$. See the study by Bell and Fishburn (2000) for an axiomatisation of functions of this type.

⁷ These properties continue to hold if $\lambda > 1$, the only difference between cases being that, if $\lambda > 1$ then, even if $u(x)$ is differentiable at $x = 0$, $v(x)$ is not differentiable at $x = 0$.

Alternatively, setting $\phi_1 = t/[ft + s]$ and $\phi_2 = 0$, we obtain the taxpayer’s post-tax income if s/he does not evade (the legal post-tax income): $R = Y [1 - t]$. This specification for reference income was first proposed by [Dhimi and al-Nowaihi \(2007\)](#). It is also considered by [Trotin \(2012\)](#), and is one of the examples in the study of [Hashimzade et al. \(2013\)](#). [Kőszegi and Rabin \(2006\)](#) make a general argument that the reference level should reflect the expected outcome of the lottery. Consistent with such approaches, reference income is equated to the expected value of the gamble by setting $\phi_1 = w(p)$ and $\phi_2 = 0$. [Hashimzade et al. \(2013\)](#) also consider the specification $R = X [1 - t]$ as one of their examples, which obtains from A5 with $\phi_1 = 1/[ft + s]$, and $\phi_2 = 0$. The specification of [Yaniv \(1999\)](#) exercises fully the generality of A5, for it implies $\phi_2 \neq 0$. The as yet undefined notation in [Table 1](#) required to characterise Yaniv’s specification we shall detail later.

Taxpayers are assumed to choose X to maximise

$$V = w(p) v(\Delta Y^c) + [1 - w(p)] v(\Delta Y^n), \tag{3}$$

where, under A5,

$$\Delta Y^n = [ft + s][\phi_1 [Y - X] + \phi_2]; \quad \Delta Y^c = [ft + s][[\phi_1 - 1][Y - X] + \phi_2].$$

[Table 2](#) introduces four variants of the TBM, each of which may be considered with or without social stigma, and under either $p' = 0$ or $p' < 0$. These are only four of many possible variants, but are sufficient to make our key observations.

The simplest variant is EUT. To understand the marginal contribution of loss aversion and probability weighting, these features are introduced in the LA \wedge PW variant. We introduce these together as it shall transpire that neither is important for the pre-

Table 1 Characterising specifications of reference income in the literature

R	ϕ_1	ϕ_2
Constant (= c)	$\frac{t}{ft+s}$	$\frac{Y[1-t]-c}{ft+s}$
$Y [1 - t]$	$\frac{t}{ft+s}$	0
$X [1 - t]$	$\frac{1}{ft+s}$	0
Exp. val.	$w(p)$	0
Yaniv (1999)	$\frac{t}{ft+s}$	$\frac{t[\omega b - Y]}{ft+s}$

Table 2 Variants of the TBM

Variant	R	Diminishing sensitivity	Loss aversion	Probability weighting
EUT	0	No	No	No
LA \wedge PW	0	No	Yes	Yes
RD	$\in (Y^c, Y^n)$	No	Yes	Yes
PT	$\in (Y^c, Y^n)$	Yes	Yes	Yes

dictions of the TBM with respect to the Puzzle. Reference dependence is introduced in variant RD, while the last variant, PT, satisfies the assumptions of cumulative prospect theory (Tversky and Kahneman 1992). As is widely noted, under diminishing sensitivity an interior maximum to (3) must satisfy $\Delta Y^n > 0$ for p'' sufficiently small, for otherwise the taxpayer’s objective function is globally convex. Moreover, if $\Delta Y^c > 0$, then the predictions of the PT variant are identical to those of the RD variant. Hence, under PT we assume $\Delta Y^c < 0 < \Delta Y^n$, and we retain this assumption in the RD variant to isolate the marginal effect of allowing for diminishing sensitivity. As the most parsimonious of the variants, EUT is at an inherent disadvantage relative to the remaining variants in respect of predicting empirical phenomena. This observation would make any failure of PT to outperform EUT the more surprising. As we shall now go on to demonstrate, these four variants of the TBM are consistent with special cases of A5.

Differentiating with respect to X , we obtain the first and second derivatives of (3) as

$$\frac{\partial V}{\partial X} = [ft + s] [w(p) [1 - \phi_1] v'(\Delta Y^c) - [1 - w(p)] \phi_1 v'(\Delta Y^n)] - p'w' [v(\Delta Y^n) - v(\Delta Y^c)]; \tag{4}$$

$$\frac{\partial^2 V}{[\partial X]^2} \equiv D = [ft + s]^2 [w(p) [1 - \phi_1]^2 v''(\Delta Y^c) + [1 - w(p)] [\phi_1]^2 v''(\Delta Y^n)] + 2[ft + s] p'w' [[1 - \phi_1] v'(\Delta Y^c) + \phi_1 v'(\Delta Y^n)] - [p''w' + [p']^2 w''] [v(\Delta Y^n) - v(\Delta Y^c)]. \tag{5}$$

Sufficient conditions for a (global) interior maximum with respect to Eqs. (4) and (5) are $\partial V/\partial X = 0$ and the second derivative $D < 0$ for all X . Under diminishing sensitivity, the second-order condition ($D < 0$)—implying that the objective function is concave in X —cannot be guaranteed for any easily interpretable restriction on the parameters. Moreover, under diminishing sensitivity it is possible—because of the possibility of corner solutions—that the first- and second-order conditions do not describe the solution of the maximisation problem.⁸ As these difficulties of PT are well understood, we choose to set them aside here. Henceforth, when analysing the TBM under diminishing sensitivity, we proceed under the maintained assumption that the second derivative in (5) is negative, such that indeed the first-order condition describes a unique choice for the taxpayer on the interval $X \in (0, Y)$.⁹

A6. $D < 0$ for all X .

If ϕ_1 is decreasing in stigma ($\phi_{1s} < 0$), as is the case for all the specifications of reference income in Table 1, then the restriction to interior solutions for evasion places

⁸ Local maxima may also arise, so the first-order condition may not possess a unique solution. See the study by Hashimzade et al. (2013) for a detailed discussion of these difficulties.

⁹ We do not investigate the properties of the corner solutions $X \in \{0, Y\}$, for the descriptive validity of tax evasion as an all-or-nothing activity appears weak.

an upper bound $s < \bar{s}$ on the level of stigma. We prove the following Lemma, which extends Dhimi and al-Nowaihi’s Proposition 5:

Lemma 1 *If $\phi_{1s} < 0$ then, at any interior maximum of (3), it must hold that $s < \bar{s}$, where \bar{s} is the unique s such that*

$$\phi_1(\bar{s}) = \begin{cases} \lim_{X \uparrow Y} \frac{\lambda w(p(X))}{1 + [\lambda - 1]w(p(X))} & \text{if } \phi_2 = 0; \\ \lim_{X \uparrow Y} w(p(X)) & \text{otherwise.} \end{cases}$$

3 Analysis

3.1 Exogenous audit probability

In this section, we examine the four variants of the TBM given in Table 2 under one further assumption regarding the probability of audit:

A7. $p' = 0$,

which is equivalent to the assumption of random auditing. As we restrict our attention to interior maxima, and several variants in Table 2 additionally require $\Delta Y^c < 0 < \Delta Y^n$, we establish the conditions under which each of these restrictions is satisfied:

Lemma 2 *Under A0–A7,*

- (i) *at any interior maximum of (3), it must hold that $\phi_1 \in (0, 1)$;*
- (ii) *$\Delta Y^c < 0 < \Delta Y^n$ always holds at an interior maximum of (3) if and only if $\phi_2 = 0$.*

Part (ii) of Lemma 2 weakens the conditions for an interior maximum to exist given in Proposition 3 of Dhimi and al-Nowaihi (2007).

Proposition 1 *If assumptions A0–A7 hold then, at an interior maximum,*

- (i) *under both EUT and $LA \wedge PW$, $\frac{\partial X}{\partial t} \geq 0$ for s sufficiently close to (or equal to) zero and $\frac{\partial X}{\partial t} < 0$ for s sufficiently close to \bar{s} ;*
- (ii) *under RD there exists a $\Phi_{(ii)} > 0$ such that $\frac{\partial X}{\partial t} \geq 0 \Leftrightarrow \phi_{1t} \leq \Phi_{(ii)}$;*
- (iii) *under PT there exists a critical value of ϕ_{1t} , $\Phi_{(iii)}$, which may be positive or negative, around which $\frac{\partial X}{\partial t}$ switches sign. For s sufficiently close to \bar{s} (or for ϕ_1 sufficiently close to zero), $\frac{\partial X}{\partial t} \geq 0 \Leftrightarrow \phi_{1t} \leq \Phi_{(iii)} > \Phi_{(ii)}$;*
- (iv) *under both PT and RD, $\frac{\partial X}{\partial t} \geq 0$ for ϕ_{1t} sufficiently close to (or equal to) zero;*
- (v) *under Yaniv’s (1999) specification of reference income there exists a $\Phi_{(v)} > -\frac{t_f}{t_f + s}$ such that $\frac{\partial X}{\partial t} \geq 0 \Leftrightarrow \varepsilon_{\phi_2, t} \leq \Phi_{(v)}$.*

Part (i) of Proposition 1 begins by extending Yitzhaki’s (1974) statement of the Puzzle—which implicitly assumes $\lambda = 1$, $s = 0$ and $w(p) = p$ —to allow for loss aversion, probability weighting and for sufficiently low levels of stigma. The finding in part (i) is a pure income effect: there is no substitution effect as both the tax rate and the penalty rate increase proportionally with t . The income effect generated by

an increase in t can itself be considered in two parts. The first part, IE_ϕ , is the part of the income effect arising from movements in R through $\{\phi_1, \phi_2\}$; and the second part, $IE_{-\phi}$, is the remainder of the effect—arising directly from movements in t and from the equilibrium adjustments of X . $IE_{-\phi}$ always increases the optimal declaration following a tax rate rise at an interior maximum, whereas IE_ϕ under both EUT and $LA \wedge PW$ is zero; hence, the Yitzhaki puzzle.

There are two elements of PT that can reverse the above intuition. Reference dependence can reverse the Puzzle if it makes taxpayers feel richer in expectation, rather than poorer, following an increase in the tax rate. To make these concepts precise, we define the expected absolute wealth of a taxpayer as $E(Y) \equiv w(p)Y^c + [1 - w(p)]Y^n$ and the expected relative wealth (when wealth is measured relative to the reference level R) as $E(\Delta Y) \equiv w(p)\Delta Y^c + [1 - w(p)]\Delta Y^n$. We write that taxpayers “feel richer in expectation” when $E(\Delta Y)$ increases.

Taxpayers can be expected to be left poorer in absolute terms from a tax rate rise, even after accounting for equilibrium adjustments in X . The only possible exception is the (seemingly unlikely) case in which the government operates at a tax rate on the “wrong” side of a Laffer curve relationship for expected revenue (whereby $\partial X/\partial t$ is sufficiently negative that also $\partial [tX]/\partial t < 0$, such that taxpayers pay less tax following a tax rate rise).¹⁰ Under EUT we have $R = 0$ so when taxpayers are made poorer in absolute terms from a tax rate rise, they are also poorer in relative terms. Yet, under reference dependence, it is possible for an increase in the tax rate to make taxpayers feel *richer* in relative terms, even when they are made poorer in absolute terms, if reference income falls faster with the tax rate than does post-tax income. When taxpayers feel richer following a tax rate rise, the income effect drives declared income down, rather than up, creating an opening to resolve the Puzzle (the RD-effect). Under what conditions will an increase in the tax rate result in expected relative income, $E(\Delta Y)$, (and relative income in the not-caught state, ΔY^n) increasing too? We now show that the answer to this question is regulated by the sign of ϕ_{1t} . In the following Lemma, we write “ d ” rather than the partial derivative symbol “ ∂ ” to emphasise that here we consider the “full” effect of a change in t , including that arising from any equilibrium adjustment in X .

Lemma 3 *Under the assumptions of Proposition 1, it holds that*

- (i) *Under RD* $\frac{d[E(\Delta Y)]}{dt} \geq 0 \Leftrightarrow \frac{d[\Delta Y^n]}{dt} \geq 0 \Leftrightarrow \phi_{1t} \geq 0$;
- (ii) *Under PT the sign of* $\frac{d[E(\Delta Y)]}{dt}$ *and* $\frac{d[\Delta Y^n]}{dt}$ *switch around* $\phi_{1t} = 0$. *For s sufficiently close to* \bar{s} , $\frac{d[E(\Delta Y)]}{dt} \geq 0 \Leftrightarrow \frac{d[\Delta Y^n]}{dt} \leq 0 \Leftrightarrow \phi_{1t} \leq 0$.

With Lemma 3 in hand, it is straightforward to interpret part (ii) of Proposition 1, which states that a necessary condition for the RD variant to reverse the Puzzle is that $\phi_{1t} > 0$. By part (i) of Lemma 3, this condition is equivalent to the statement that a tax rate rise makes taxpayers feel richer in expectation. As a sufficient condition to reverse the Puzzle, ϕ_{1t} must not just be positive, but sufficiently so to overcome an

¹⁰ Fisman and Wei (2004) empirically document an instance in which tax evasion indeed leads to a Laffer curve relationship holding. Clearly, however, were the government operating on the wrong side of a Laffer curve, it is unclear why it would ever wish to further increase the tax rate.

offsetting substitution effect. How plausible is the necessary assumption that taxpayers feel richer in the not-caught state, and in expectation, following a tax rate rise? Barring the Laffer curve consideration given above, the taxpayer—judged in absolute terms—is poorer in expectation following a tax rate increase, in which case the RD-effect relies on a strong disjunction between the response to a tax rate rise of absolute and relative income to resolve the Yitzhaki puzzle. Although we know of no direct empirical evidence on this point, we find the idea that taxpayers made poorer in absolute terms would also feel poorer (i.e., $d[E(Y)]/dt < 0 \Rightarrow d[E(\Delta Y)]/dt < 0$) compelling from a psychological perspective.

The second element of PT that can reverse the Yitzhaki puzzle is diminishing sensitivity: this can overturn the Puzzle by reversing the sign of the income effect in the caught state, so that when a taxpayer feels poorer in the caught state, they are induced to take more (not less) risk, and thus declare less income (the DS-effect). Note, however, that whereas the RD-effect requires that a tax rate rise makes taxpayers feel richer, the DS-effect requires that a tax rate rise makes taxpayers feel poorer. Equivalently, reference dependence requires reference income to be sensitive to the tax rate to resolve the Puzzle, but diminishing sensitivity requires reference income to be insensitive to the tax rate. Therefore, the effects of reference dependence and diminishing sensitivity with respect to the Puzzle oppose each other. Accordingly, part (ii) of Lemma 3 implies that an increase in ϕ_{1t} (which makes reference income more sensitive to the tax rate) can either increase or decrease the effect of a tax rate rise on expected income depending on the balance of the DS- and RD-effects. It is this ambiguity that lies at the heart of part (iii) of Proposition 1, which gives the condition needed for the DS-effect to reverse the Puzzle.

Part (iv) of the Proposition exploits the observation that the RD- and DS-effects are in opposition to establish a simple and relevant condition under which the Yitzhaki puzzle unambiguously holds. It states that when the expected value of the gamble moves sufficiently little with the tax rate (ϕ_{1t} close to zero), the Yitzhaki puzzle holds under both the PT and RD variants. In this case, taxpayers neither feel sufficiently richer after a tax rate rise for the RD-effect to overturn the Puzzle, nor feel sufficiently poorer for the DS-effect to overturn the Puzzle.

How does introducing stigma mediate the above findings? Part (i) of Proposition 1 states that the Yitzhaki puzzle is always resolved under EUT for s sufficiently high. This finding is consistent with various analyses incorporating stigma—differing in the way in which stigma enters the taxpayer's objective function—that establish bounds for s above which $\partial X/\partial t < 0$ (al-Nowaihi and Pyle 2000; Dell'Anno 2009; Gordon 1989; Kim 2003). In contrast, even in the neighbourhood of the maximum level of stigma, the Puzzle is not unambiguously reversed in either the PT or RD variants. Rather, close to the maximum level of stigma, part (iii) of Proposition 1 states that the same questionable necessary condition for reversing the Yitzhaki puzzle under the RD variant holds also under PT: taxpayers must feel richer in the not-caught state, and in expectation, after an increase in the tax rate (and, as a sufficient condition, this effect must be sufficiently strong).

3.1.1 Implications for the literature

We now consider the implications of our findings for the existing literature. We first consider those specifications of reference income given in Table 1 that satisfy $R \in (Y^c, Y^n)$. To begin, we consider these specifications in the absence of stigma ($s = 0$). In this case, the specification of reference income as the taxpayer's legal post-tax income, $R = Y [1 - t]$, implies $\phi_1 = f^{-1}$, and the specification as the expected value of the tax gamble implies $\phi_1 = w(p)$. It is readily observed that, for both specifications, $\phi_{1t} = 0$, thus (by part (iv) of Proposition 1) the Yitzhaki puzzle holds both under the PT and RD variants.¹¹ Only for the specification of reference income $R = X [1 - t]$ examined by Hashimzade et al. (2013) does part (iv) of Proposition 1 not apply, for it implies $\phi_1 = [ft]^{-1}$, hence $\phi_{1t} = -[ft^2]^{-1} < 0$. In this case, the Yitzhaki puzzle holds under RD, but the sign of $\partial X/\partial t$ is not determined a priori under PT. Thus, for two of the three specifications of reference income in the literature satisfying $R \in (Y^c, Y^n)$ —including the most utilised specification of reference income as the taxpayer's legal post-tax income—neither PT nor its stripped-down variants resolve the Yitzhaki puzzle. Note that, for these two specifications that do not resolve the Puzzle, the ability of the PT and RD variants to reverse the Yitzhaki puzzle is strictly weaker than that of EUT. The latter can always reverse the Puzzle, albeit by invoking the empirically unsatisfactory assumption of increasing absolute risk aversion (and this must be sufficiently strong), whereas PT and its variants cannot reverse the Puzzle for any choice of preferences consistent with an interior maximum.

Allowing for stigma yields mixed findings. For the specification of reference income as legal post-tax income, the predictions of the PT and RD variants are improved, for we obtain $\phi_{1t} = s/[ft + s]^2 > 0$, such that the sign of $\partial X/\partial t$ becomes ambiguous in general, but unambiguously negative for s close to \bar{s} . This reversal as $s \uparrow \bar{s}$ relies, however, on a tax rate rise making taxpayers feel richer ($\phi_{1t} > 0$). The predictions for $\partial X/\partial t$ are unchanged for the specification of reference income as the expected value of the gamble, however, and for the specification $R = X [1 - t]$ they worsen: in the absence of stigma this specification implies an ambiguous sign for $\partial X/\partial t$ under PT, but unambiguously implies Yitzhaki's puzzle for s close to \bar{s} . In summary, augmenting EUT with stigma unambiguously improves its descriptive abilities, for it always overturns the Yitzhaki puzzle if enough stigma is allowed for. Augmenting PT with stigma sharpens its predictions with respect to the Yitzhaki puzzle, but does not unambiguously improve their descriptive validity. In particular, for either the PT or RD variants to overturn the Puzzle, even close to the maximum level of stigma, an increase in the tax rate must make taxpayers feel richer in expectation.

Proposition 1 connects to two further studies in the existing literature: Yaniv (1999) and Bernasconi and Zanardi (2004), each of which employs a specification of reference income that does not always satisfy $R \in (Y^c, Y^n)$. Specifically, Yaniv examines a model with reference income specified as $R = Y - H$, where H is the amount of

¹¹ Trotin (2012) uses the taxpayer's legal post-tax income to specify reference income in a model without stigma and claims (her Proposition 8) to resolve the Yitzhaki puzzle. The difference between our findings and hers is attributable to a non-sequitur in her proof of Proposition 8. In particular, we are unable to replicate the expression for $\partial \Phi_{\bar{R}}(x^*, t)/\partial t$ in the first line of her proof.

an advance tax payment. The advance payment H is specified (up to a constant) as $H = \omega t b$ where b is the tax authority's estimate of the taxpayer's income (which could under- or over-estimate the true Y), and $\omega \in [0, 1]$. This specification of reference income is a special case of A5 with $\phi_1 = t/[tf + s]$ and $\phi_2 = t[\omega b - Y]/[ft + s]$. Yaniv's model is not, however, encompassed by the TBM for, in forming the taxpayer's objective function, the author adopts the segregation assumption of the original 1979 version of PT, according to which certain gains and losses are extracted from the gamble. The coding phase (in which the original version of PT held segregation to occur) is, however, de-emphasised in the later cumulative version of PT—the version we employ. Accordingly, the remaining literature on PT and tax evasion, all of which post-dates Yaniv's contribution, does not adopt this assumption. Does Yaniv's model still reverse the Yitzhaki puzzle if segregation is not assumed? In part (v) of Proposition 1, we give a result that addresses this question, and answers it in the negative: for Yaniv's specification of R we obtain $\varepsilon_{\phi_2,t} = -1$ yet, by part (v), a necessary (and still not sufficient) condition for the Yitzhaki puzzle to be overturned is $\varepsilon_{\phi_2,t} > -tf/[tf + s] \geq -1$.

The last contribution to the literature we discuss here is that of [Bernasconi and Zanardi \(2004\)](#). Unlike the rest of the literature, this study does not explicitly specify reference income, but rather it examines taxpayer behaviour for all possible values of R . Implicitly, however, the non-specification of R is equivalent to the specification of R as a constant, $R = c$. Viewed this way, the Bernasconi and Zanardi specification of R is the special case of A5 with $\phi_1 = t/[ft + s]$ and $\phi_2 = [Y[1 - t] - c]/[ft + s]$. As this specification of reference income implies $R_t = 0$ the RD-effect does not arise, leaving only a pure DS-effect. Hence, under the PT variant, $\partial X/\partial t < 0$. The contribution of Proposition 1 in respect of this analysis is to highlight that, once an endogenous specification for R is adopted, PT and its variants may no longer reverse the Yitzhaki puzzle. As the authors note in their conclusion, pinning-down reference income is unavoidable if PT is to yield clear and testable predictions.

We complete our discussion of Proposition 1 by assessing its implications for the separate elements of the TBM. The proposition makes clear that the Yitzhaki puzzle can be resolved without recourse to either probability weighting or loss aversion, and part (i) makes clear that when EUT is augmented with probability weighting and/or loss aversion the Puzzle is unaffected. These two features of PT are, therefore, unimportant with respect to the predicted sign of $\partial X/\partial t$.¹² Diminishing sensitivity, too, is neither necessary nor sufficient to resolve the Puzzle. Indeed, in the presence of social stigma, diminishing sensitivity appears to hinder the resolution of the Puzzle. To see this, note from part (iv) of Proposition 1 that, close to the maximum level of social stigma, the necessary condition for the Puzzle to be resolved is weaker under RD than under PT. Thus, reference dependence stands out as the only feature of PT that is essential to resolving the Puzzle.

¹² Consistent with our finding regarding probability weighting, [Eide \(2001\)](#) shows that introducing (rank-dependent) probability weighting into the standard tax evasion model does not qualitatively change the comparative statics results.

3.2 Endogenous audit probability

In practice only a fairly small proportion of tax authority audits are selected randomly. Accordingly, here we repeat the analysis conducted in the previous section under an alternative assumption to that of A7:

A8. $p' < 0$.

Under A8, [Dhmi and al-Nowaihi \(2007\)](#) prove (their Proposition 1c) that if the only assumptions made with regard to preferences are A0, A2 and A3, then the sign of $\partial X/\partial t$ under EUT is ambiguous. To obtain comparative static results under A8, we require a stronger restriction on preferences. Homogeneous utility appears of particular relevance in this context as Tversky and Kahneman’s (1992) original value function—subsequently axiomatised under PT by [al-Nowaihi et al. \(2008\)](#)—is piecewise homogeneous. Accordingly, we assume

$$A9. \quad xu'(x) = \begin{cases} \beta u(x) & \text{if } x \geq 0; \\ \gamma u(x) & \text{otherwise;} \end{cases}$$

which is equivalent to the statement that $u(x)$ is homogeneous of degree β in the gain domain, and of degree γ in the loss domain.¹³ For A9 to be compatible with A2, we require $\beta \in (0, 1)$ and $\gamma > 1$.

Whereas, when $p' = 0$, $\phi_1 \in (0, 1)$ was guaranteed by Lemma 2, under A8, interior maxima satisfying $\phi_1 > 1$ can exist. This is readily observed from the first-order condition in equation (4), where the final term is positive when $p' < 0$. We now show, however, that under the stronger preference restriction in A9 $\phi_1 \in (0, 1)$ must hold at an interior maximum when $\phi_2 = 0$.

Lemma 4 *Under A0–A6 and A8–A9, and assuming $\phi_2 = 0$, at any interior maximum of (3), it must hold that $\phi_1 \in (0, 1)$.*

Given Lemma 4, it is straightforward to verify that $\phi_2 = 0$ remains necessary and sufficient for $\Delta Y^c < 0 < \Delta Y^n$ to hold at any interior maximum.

Proposition 2 *If assumptions A0–A6 and A8–A9 hold then, at an interior maximum,*

- (i) *under both EUT and $LA \wedge PW$ $\frac{\partial X}{\partial t} \geq 0$ for s sufficiently close to (or equal to) zero, and $\frac{\partial X}{\partial t} < 0$ for s sufficiently close to \bar{s} ;*
- (ii) *under PT $\frac{\partial X}{\partial t} \geq 0 \Leftrightarrow \phi_{1t} \leq 0$;*
- (iii) *under RD $\frac{\partial X}{\partial t} \geq 0 \Leftrightarrow \phi_{1t} \leq \frac{f[\gamma-\beta]\phi_1[\phi_1-1]}{[tf+s][\gamma\phi_1-\beta[\phi_1-1]]} < 0$.*

Part (i) of Proposition 2 extends the original statement of the Yitzhaki puzzle to the case of variable audit probability and homogeneous utility (as well as allowing for loss aversion, probability weighting, and sufficiently low levels of stigma). Introducing variable audit probabilities does not, therefore, appear to enhance importantly the

¹³ An example $u(x)$ that satisfies A0–A4 is given by the piecewise function $u(x) = ax^\beta$ for $x \geq 0$ and $u(x) = ax^{1/\beta}$ for $x < 0$, where $a > 0$ and $\beta \in (0, 1)$.

descriptive abilities of EUT with respect to the Puzzle. Under PT, the income and substitution effects cancel, leaving a residual that takes the sign of $-\phi_{1t}$. A simple necessary and sufficient condition for the Yitzhaki puzzle to be resolved, therefore, holds: $\phi_{1t} > 0$. In the absence of stigma, all three of the specifications of reference income satisfying $R \in (Y^c, Y^n)$ in the existing literature fail to satisfy this necessary condition. Moreover, it is straightforward to show (with the same procedure we use to prove Lemma 3) that when $\phi_{1t} > 0$ it holds that both $d[\mathbf{E}(\Delta Y)]/dt$ and $d[\Delta Y^n]/dt$ are positive, so whenever the Puzzle is resolved under PT it also predicts that expected relative wealth increases following a tax rate rise. Thus, allowing for variable audit probabilities does not appear to enhance the descriptive abilities of PT and its variants with respect to the Puzzle either. Once again, the condition for the RD variant to overturn the Yitzhaki puzzle is strictly weaker than that for PT.

The effects of stigma are similar to those in the analysis with fixed probabilities. In particular, in the presence of stigma, the specification of reference income as the legal post-tax income satisfies the necessary condition for Yitzhaki's puzzle to be resolved, but the predictions regarding $\partial X/\partial t$ of the PT and RD variants for the remaining two specifications of reference income satisfying $R \in (Y^c, Y^n)$ are either unchanged with stigma, or worsen with stigma.

What are the implications of Proposition 2 for the existing literature? The model of [Dhami and al-Nowaihi \(2007\)](#) satisfies the assumptions of Proposition 2, where it corresponds to a special case of the PT variant with $R = Y[1 - t]$. Our insights with respect to this contribution are twofold. First, while these authors demonstrate, in this special case, that PT outperforms EUT, we can assess whether PT has a general advantage over EUT in relation to predicting the sign of $\partial X/\partial t$. Concerning this question, Proposition 2 does not suggest that PT has such an advantage.¹⁴ Second, the specification of the model implies that taxpayers who are made poorer in absolute terms by a tax rate rise will feel relatively richer.

4 Conclusion

Prospect theory (PT) is widely viewed as the best available description of how (many) people behave in risky settings, and we do not dissent from this view. [Barberis \(2013\)](#) notes, however, that PT is not always straightforward to apply: in particular, the most appropriate specification of the reference level is often unclear.¹⁵ Yet, we show that theoretical predictions related to tax evasion and the Yitzhaki puzzle depend crucially on the specification of that reference level.

¹⁴ We have shown (Table 1) that the specification of reference income in the study by [Dhami and al-Nowaihi \(2007\)](#) corresponds to a convex combination of Y^c and Y^n with a weight on Y^c of $\alpha = \phi_1 = t/[ft + s]$. A possible confusion in evaluating [Dhami and al-Nowaihi's](#) result, which our analysis enables us to clarify, is that these authors assert this specification to be the only one with the property $R \in (Y^c, Y^n)$. We show, however, that this specification is just one special case of a wider set of specifications with this property. In particular, any specification of reference income as a (strictly) convex combination of Y^c and Y^n must also satisfy this property (Lemma 2).

¹⁵ The specification of the value and weighting functions of PT may also be problematic (see, e.g., [Neilson and Stowe 2002](#); [Sadiraj 2014](#)).

We focus on tax evasion and in particular on the Yitzhaki puzzle: the EUT model of tax evasion predicts a decrease in tax evasion when the tax rate increases ($\partial X/\partial t < 0$). The PT literature provides several specific examples in which the Puzzle is solved: our research is a first step towards understanding, in a general setting, the conditions under which PT reverses the Puzzle. Furthermore, by stripping down the elements of PT, we are able to (i) identify the forces that drive predictions concerning the impact of the tax rate on tax evasion; (ii) compare our findings with those under Expected Utility Theory (EUT). Although our findings do not clearly endorse the descriptive validity of EUT, we find that, in respect of the Puzzle, many specifications of PT share similar descriptive deficiencies. Interestingly, the set of specifications of PT that cannot reverse the Puzzle overlaps with many of those advocated in the literature. The auxiliary assumptions we consider—social stigma and variable audit probability—allow PT in some cases to reverse the Yitzhaki puzzle, but when they do so, they also allow the model under EUT to do likewise when similarly augmented. Although there are indeed many specifications of PT that do reverse the Puzzle, we find that in many instances these specifications imply that taxpayers will feel richer after a tax rise, for they imply that reference income falls (and by a sufficient amount) following a tax rate increase. That taxpayers made objectively poorer by a tax rate rise would feel subjectively richer seems psychologically questionable, albeit empirical work addressing this point is needed. Thus, Barberis's point concerning the difficulty of properly identifying the reference level is evident in the tax evasion context.

Our findings contribute to the literature on the individual elements of PT. Of the four elements of PT that we isolate, only reference dependence is necessary to resolve the Puzzle, which is in agreement with its supposed status as the most accepted of all the elements of PT. Loss aversion and probability weighting are largely irrelevant for the sign of $\partial X/\partial t$. Invoking Occam's razor, we believe that some of the findings relating to the Yitzhaki puzzle that have been attributed to PT may more properly be interpreted as being attributable to simpler reference-dependent models that contain only a subset of the elements of PT. Perhaps more surprisingly, diminishing sensitivity is neither necessary nor sufficient to resolve the Yitzhaki puzzle, and indeed pushes $\partial X/\partial t$ the "wrong" way in the presence of social stigma.

Given the importance of PT in explaining behaviour in so many economic domains, it seems altogether likely that it is also of importance in explaining behaviour towards tax evasion too. As such, our cautionary findings need not imply that tax evasion researchers should cease to explore the insights of PT, but they do suggest that renewed focus must be placed on the specification of reference income and preferences. In particular, if researchers wish to retain the common assumption of homogeneous preferences then specifications of reference income beyond those encompassed by our model must be sought, for otherwise taxpayers must feel subjectively richer after a tax rate rise for PT to reverse the Puzzle.

Acknowledgements Rablen acknowledges financial support from the ESRC (ES/K001744/1), and Piolatto from the Spanish Ministry of Science and Innovation (ECO-2015-68311-R) and the Generalitat de Catalunya (2014SGR420). We thank the editor; two anonymous referees; participants of PET13 (Lisbon) and PEUK2014 (Bristol); seminar audiences at Durham, Exeter, Haverford, Kent, Lucca, Milan Bicocca,

and Sheffield; Ali al-Nowaihi and Sanjit Dhmi for many useful comments on a preliminary version of the manuscript.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix

Proof of Lemma 1 Under A6 the first derivative in (4) is decreasing, and when an interior maximum exists it must also switch sign on $(0, Y)$. Hence, $\lim_{X \downarrow 0} \partial V / \partial X \geq 0$ and $\lim_{X \uparrow Y} \partial V / \partial X \leq 0$. Taking the limit of (4) as $X \uparrow Y$, and noting that $\lim_{X \uparrow Y} \{v(\Delta Y^n) - v(\Delta Y^c)\} = 0$ by A1, the latter inequality in the preceding sentence implies

$$\phi_1 \geq \lim_{X \uparrow Y} \frac{w(p(Y)) v'(\Delta Y^c)}{[1 - w(p(Y))] v'(\Delta Y^n) + w(p(Y)) v'(\Delta Y^c)}.$$

Using (1), this may be re-written as

$$\phi_1 \geq \begin{cases} \lim_{X \uparrow Y} \frac{\lambda w(p(X))}{1 - w(p(X)) + \lambda w(p(X))} & \text{if } \phi_2 = 0; \\ \lim_{X \uparrow Y} w(p(X)) & \text{otherwise.} \end{cases}$$

□

Proof of Lemma 2 (i) Suppose, by way of contradiction, $\phi_1 \geq 1$ then the first derivative in (4) with $p' = 0$ is always negative, thus ruling out the possibility of an interior maximum. Hence, $\phi_1 < 1$ at an interior maximum. Next, the proof of Lemma 1 establishes that either $\phi_1 \geq \lim_{X \uparrow Y} w(p(X)) > 0$ or

$$\phi_1 \geq \lim_{X \uparrow Y} \frac{\lambda w(p(X))}{1 - w(p(X)) + \lambda w(p(X))} \geq w(p(X)) > 0.$$

Hence, $\phi_1 > 0$.

(ii) We first show that $\phi_2 = 0$ implies that $Y^c < R < Y^n$. Equation (2) becomes $R = \phi_1 Y^c + [1 - \phi_1] Y^n$. By Lemma 2, $\phi_1 \in (0, 1)$, hence R is a convex combination of Y^n and Y^c .

We now show that $Y^c < R < Y^n$ for all $X \in (0, Y)$ implies $\phi_2 = 0$. $\Delta Y^n > 0 > \Delta Y^c$ holds for all $X \in (0, Y)$. At $X = Y$, however, it holds that $Y^n = Y^c = Y[1 - t]$, so $\Delta Y^n = \Delta Y^c = 0$. Therefore, as $R \in [Y^n, Y^c]$ arbitrarily close to $X = Y$, by continuity, we must have $R = Y^n = Y^c = Y[1 - t]$ at $X = Y$. However, when $X = Y$, Eq. (2) gives $R = Y[1 - t] - [ft + s]\phi_2$, which is consistent with $R = Y[1 - t]$ also holding only if $[ft + s]\phi_2 = 0$, hence $\phi_2 = 0$. □

Proof of Proposition 1 We start by proving statement (i). Under the parameter restrictions defining $LA \wedge PW$ the derivative $\partial X/\partial t$, found implicitly from (4), is

$$\frac{\partial X}{\partial t} = -\frac{1}{D} \left[[1-w(p)]tXv''(Y^n) - w(p)\{t[f-1] + s\}[X+f[Y-X]]v''(Y^c) + w(p)[f-1]v'(Y^c) - [1-w(p)]v'(Y^n) \right].$$

Substituting from the first-order condition, we obtain

$$\frac{\partial X}{\partial t} = -\frac{1}{D} \left\{ [1-w(p)]tv'(Y^n)[Y[A(Y^c) - A(Y^n)] + [Y-X][A(Y^n) + [f-1]A(Y^c)]] - \frac{w(p)s}{t}v'(Y^c) \right\}.$$

If $s = 0$ this rewrites as

$$\frac{\partial X}{\partial t} = \frac{1}{t} \left[[Y-X] + \frac{Y[A(Y^c) - A(Y^n)]}{[f-1]A(Y^c) + A(Y^n)} \right]. \tag{A.1}$$

Under A4, $A(Y^c) - A(Y^n) \geq 0$, hence $\partial X/\partial t > 0$. As equation (A.1) holds for $w(p) = p$, and is independent of λ , $\partial X/\partial t \geq 0$ also holds under EUT. If $s \uparrow \bar{s}$ then $X \uparrow Y$ and $A(Y^c) \downarrow A(Y^n)$ so

$$\frac{\partial X}{\partial t} = \lim_{s \uparrow \bar{s}} \frac{w(p)s}{tD} v'(Y^c) < 0.$$

(ii–iv). Under the relevant parameter restrictions the derivative $\partial X/\partial t$, found implicitly from the first-order condition, is

$$\frac{\partial X}{\partial t} = \frac{[ft+s]}{D} \times \left[\begin{aligned} &\phi_{1t} [w(p)v'(\Delta Y^c) + [1-w(p)]v'(\Delta Y^n)] \\ &+ [Y-X] \left[\begin{aligned} &w(p)[\phi_1-1][f[\phi_1-1] + \phi_{1t}[ft+s]]v''(\Delta Y^c) \\ &+ \phi_1[1-w(p)][f\phi_1 + \phi_{1t}[ft+s]]v''(\Delta Y^n) \end{aligned} \right] \end{aligned} \right].$$

After some algebra and using the first-order condition to simplify the equation, we obtain:

$$\frac{\partial X}{\partial t} = \frac{f[Y-X]}{ft+s} + \frac{[ft+s]\phi_{1t}}{D} \left[\frac{1}{\phi_1} [w(p)v'(\Delta Y^c)] + \phi_1[ft+s][Y-X][1-w(p)]v'(\Delta Y^n)[A(\Delta Y^c) - A(\Delta Y^n)] \right]. \tag{A.2}$$

Equation (A.2) allows us to prove parts (ii–iv). Notice that the first term in (A.2) is always positive. The second term is all multiplied by ϕ_{1t} . This is sufficient to prove

statement (iv). Indeed, when $\phi_{1t} = 0$, then the second term disappears and $\partial X/\partial t \geq 0$. By continuity, this result must also hold in a neighbourhood of 0.

Under RD, $A(\Delta Y^c) - A(\Delta Y^n) \geq 0$, so the sign of the second term in (A.2) is that of $-\phi_{1t}$. Define $\Phi_{(ii)}$ as the value of ϕ_{1t} such that $\partial X/\partial t = 0$. Then,

$$\Phi_{(ii)} = -\frac{f[Y - X]D}{[ft + s]^2} \times \left[\frac{w(p)v'(\Delta Y^c)}{\phi_1} + \phi_1[ft + s][Y - X][1 - w(p)]v'(\Delta Y^n)[A(\Delta Y^c) - A(\Delta Y^n)] \right]^{-1}.$$

Notice that $\Phi_{(ii)} > 0$. Statement (ii) follows immediately.

To prove part (iii) notice that $A(\Delta Y^c) - A(\Delta Y^n) < 0$ under PT. Then, although $\Phi_{(iii)}$ is written identically to $\Phi_{(ii)}$, its sign is no longer determined. If the term in square brackets is positive, it follows that $\partial X/\partial t \geq 0$ if $\phi_{1t} \leq \Phi_{(iii)}$ and $\partial X/\partial t < 0$ if $\phi_{1t} > \Phi_{(iii)}$. However, if the term in brackets is negative, then $\partial X/\partial t \geq 0$ if $\phi_{1t} > \Phi_{(iii)}$ and $\partial X/\partial t < 0$ if $\phi_{1t} < \Phi_{(iii)}$. Finally, notice that the term in square brackets is positive if its first (positive) term outweighs the second. This occurs when $\phi_1 = 0$ (and by continuity, in its neighbourhood) and when s tends to \bar{s} (for then X tends to Y and $[A(\Delta Y^c) - A(\Delta Y^n)]$ tends to zero).

(v) Under the relevant parameter restrictions, the derivative $\partial X/\partial t$ is

$$\frac{\partial X}{\partial t} = \frac{ft + s}{D} \left[[f[[\phi_1 - 1][Y - X] + \phi_2] + [ft + s]\phi_{2t}][\phi_1 - 1]w(p)v''(\Delta Y^c) + [f[\phi_1[Y - X] + \phi_2] + [ft + s]\phi_{2t}]\phi_1[1 - w(p)]v''(\Delta Y^n) \right].$$

Using the first-order condition in (4) and replacing $\varepsilon_{\phi_{2,t}} \equiv t[\phi_{2t}/\phi_2]$, we obtain

$$\frac{\partial X}{\partial t} = \frac{f[Y - X]}{ft + s} + \frac{\phi_2}{t} [tf + [ft + s]\varepsilon_{\phi_{2,t}}] \times \frac{A(\Delta Y^c) - A(\Delta Y^n)}{[ft + s][[\phi_1 - 1]A(\Delta Y^c) - \phi_1A(\Delta Y^n)]}.$$

The second-order condition in (5) writes as

$$\frac{\partial^2 V}{[\partial X]^2} = [ft + s]^2 \left[w(p)[1 - \phi_1]^2 v''(\Delta Y^c) + [1 - w(p)][\phi_1]^2 v''(\Delta Y^n) \right] < 0, \tag{A.3}$$

hence

$$[\phi_1 - 1]A(\Delta Y^c) - \phi_1A(\Delta Y^n) < 0.$$

As also $A(\Delta Y^c) - A(\Delta Y^n) < 0$ (under diminishing sensitivity), a sufficient condition for $\partial X/\partial t > 0$ is, therefore, $\varepsilon_{\phi_{2,t}} < -ft/[ft + s]$. □

Proof of Lemma 3 We first prove the statements relative to $d[\mathbf{E}(\Delta Y)]/dt$ and then move on to those relative to $d[\Delta Y^n]/dt$. From $\mathbf{E}(\Delta Y) = [ft + s][Y - X][\phi_1 - w(p)]$ we compute

$$\begin{aligned} \frac{d\mathbf{E}(\Delta Y)}{dt} &= f[Y - X][\phi_1 - w(p)] - [ft + s]X_t[\phi_1 - w(p)] \\ &\quad + [ft + s][Y - X]\phi_{1t} = [\phi_1 - w(p)][ft + s] \\ &\quad \times \left[\frac{f[Y - X]}{[ft + s]} - X_t \right] + [ft + s][Y - X]\phi_{1t}. \end{aligned}$$

Replacing X_t by its expression, we obtain $d\mathbf{E}(\Delta Y)/dt = \Omega\phi_{1t}$ with

$$\begin{aligned} \Omega &\equiv \frac{-1}{[\phi_1 - 1]A(\Delta Y^c) - \phi_1 A(\Delta Y^n)} \\ &\quad \times \left[[Y - X][ft + s] \left[\frac{[1 - w(p)]A(\Delta Y^c)}{+w(p)A(\Delta Y^n)} \right] + \frac{[\phi_1 - w(p)][pv'(\Delta Y^c)]}{[1 - p][\phi_1]^2 v'(\Delta Y^n)} \right]. \end{aligned}$$

Studying the sign of Ω , by the second-order condition in (A.3), the initial fraction is positive. Within the main square bracket, the first term is positive under RD while the second term is always positive, for $\phi_1 > w(p)$ by Lemma 2. This proves that under RD $d[\mathbf{E}(\Delta Y)]/dt \geq 0 \Leftrightarrow \phi_{1t} \geq 0$. Under PT, the first term in brackets may take negative values. However, for $s = \bar{s}$ (and, by continuity, in its neighbourhood), then $Y = X$ and the first term in square bracket is zero. This proves that for s sufficiently close to \bar{s} , $d[\mathbf{E}(\Delta Y)]/dt \geq 0 \Leftrightarrow \phi_{1t} \geq 0$.

Moving to ΔY^n , from $\Delta Y^n = \phi_1 [ft + s][Y - X]$ we compute

$$\begin{aligned} \frac{d[\Delta Y^n]}{\partial t} &= [f\phi_1 + [ft + s]\phi_{1t}][Y - X] - \phi_1 [ft + s]X_t \\ &= [ft + s] \left[[Y - X] \left[1 - \frac{[ft + s]^2 [\phi_1]^2 [1 - w(p)] v'(\Delta Y^n) [A(\Delta Y^c) - A(\Delta Y^n)]}{D} \right] \right. \\ &\quad \left. - \frac{[ft + s]}{D} [w(p) v'(\Delta Y^c)] \right] \phi_{1t}. \end{aligned}$$

Under RD, $A(\Delta Y^c) - A(\Delta Y^n) > 0$; hence, all the terms in the bracket are positive. This directly implies that $d[\Delta Y^n]/dt \geq 0 \Leftrightarrow \phi_{1t} \geq 0$. Instead, under PT we have $A(\Delta Y^c) - A(\Delta Y^n) < 0$. However, for $s = \bar{s}$ (and, by continuity, in its neighbourhood), then $Y = X$ and the first term in square bracket is zero: $\partial d[\Delta Y^n]/\partial t = -\frac{[ft + \bar{s}]^2}{D} [pv'(\Delta Y^c)]\phi_{1t}$. This proves that for s sufficiently close to \bar{s} , $d[\Delta Y^n]/dt \geq 0 \Leftrightarrow \phi_{1t} \geq 0$. \square

Proof of Lemma 4 Setting $\phi_2 = 0$ the taxpayer’s objective function can be written as

$$\begin{aligned} V &= w(p) [[ft + s][Y - X]]^{\beta^{1-DS}\gamma^{DS}} v(\phi_1 - 1) \\ &\quad + [1 - w(p)][ft + s]^\beta [Y - X]^\beta v(\phi_1). \end{aligned} \tag{A.4}$$

Thus, $\partial V/\partial X$ is found as

$$\begin{aligned}
 & - [tf + s]^{\beta^{1-DS}\gamma^{DS}} [Y - X]^{\beta^{1-DS}\gamma^{DS}-1} \\
 & \times \left[\begin{aligned} & [\beta^{1-DS}\gamma^{DS} w(p) - [Y - X] w' p'] v(\phi_1 - 1) \\ & + [\beta^{1-DS}\gamma^{DS} [1 - w(p)] + [Y - X] w' p'] v(\phi_1) \end{aligned} \right].
 \end{aligned}$$

If $\phi_1 \geq 1$ then $\partial V/\partial X < 0$ and if $\phi_1 \leq 0$ then $\partial V/\partial X > 0$. In either case, there cannot be an interior maximum. Hence, at an interior maximum $\phi_1 \in (0, 1)$. \square

Proof of Proposition 2 (i) Under LA \wedge PW the first-order condition is written, using A9, as

$$\begin{aligned}
 \frac{\partial V}{\partial X} &= w(p) \{t[f - 1] + s\} v'(Y^c) - [1 - w(p)] t v'(Y^n) \\
 &+ w' p' [v(Y^c) - v(Y^n)] = 0.
 \end{aligned} \tag{A.5}$$

Thus, $\partial X/\partial t$ is found as

$$\frac{\partial X}{\partial t} = -\frac{1}{\beta D} \left\{ \begin{aligned} & [\beta w(p) [f - 1] - w' p' [X + f [Y - X]] v'(Y^c) \\ & - [\beta [1 - w(p)] - w' p' X] v'(Y^n) \\ & - [\beta w(p) \{t[f - 1] + s\} + w' p' Y^c] [X + f [Y - X]] v''(Y^c) \\ & + [\beta t [1 - w(p)] + w' p' Y^n] X v''(Y^n) \end{aligned} \right\}.$$

If $s = 0$, and noting that $A(Y^i) = [1 - \beta]/Y^i, i = c, n, \partial X/\partial t$ reduces to

$$\frac{\partial X}{\partial t} = -\frac{Y}{\beta t D} \left\{ \begin{aligned} & w' p' [v'(Y^n) - v'(Y^c) + [1 - \beta] \left[\frac{Y^n - Y^c}{Y^n} \right] v'(Y^c)] \\ & + \beta t w(p) [f - 1] [1 - \beta] \left[\frac{Y^n - Y^c}{Y^n Y^c} \right] v'(Y^c) \end{aligned} \right\}. \tag{A.6}$$

Now note that from the first-order condition and A9 we have

$$\frac{\beta t w(p) [f - 1] + w' p' Y^c}{\beta t [1 - w(p)] + w' p' Y^n} = \frac{v'(Y^n)}{v'(Y^c)} = \left[\frac{Y^c}{Y^n} \right]^{1-\beta}.$$

Then, defining $\phi(z) = z^{1-\beta}$, the first-order Taylor series expansion of $\phi(z)$ at $z = 1$ is $g(z) = 1 - [1 - \beta][1 - z]$. Clearly, $g(z) \geq \phi(z)$ for all z and $g(z) > z^{1-\beta}$ for all $z \neq 1$, so

$$\frac{v'(Y^n)}{v'(Y^c)} < g\left(\frac{Y^c}{Y^n}\right) = 1 - [1 - \beta] \left[\frac{Y^n - Y^c}{Y^n} \right].$$

Rearranging, we obtain

$$v'(Y^n) - v'(Y^c) + [1 - \beta] \left[\frac{Y^n - Y^c}{Y^n} \right] v'(Y^c) < 0, \tag{A.7}$$

which signs (A.6) positive. As Eq. (A.6) holds for $w(p) = p$, and it is independent of λ , $\partial X/\partial t \geq 0$ also holds under EUT. If $s \uparrow \bar{s}$ then $\partial X/\partial t$ instead writes as

$$\frac{\partial X}{\partial t} = \lim_{s \uparrow \bar{s}} \frac{w(p)s}{tD} v'(Y^c) < 0.$$

(ii)–(iii) Setting $\phi_2 = 0$ the taxpayer’s objective function is given by (A.4). Using the first order condition, we obtain

$$\begin{aligned} \frac{\partial X}{\partial t} &= \frac{[tf + s]^{\beta^{1-DS}\gamma^{DS}} [Y - X]^{\beta^{1-DS}\gamma^{DS}-1}}{D} \\ &\times \left[\beta^{1-DS}\gamma^{DS} w(p) - [Y - X] w'p' \right] v(\phi_1 - 1) \\ &\times \left[f \left[\beta^{1-DS}\gamma^{DS} - \beta \right] + \phi_{1t} [ft + s] \frac{\beta^{1-DS}\gamma^{DS}\phi_1 - \beta [\phi_1 - 1]}{\phi_1 [\phi_1 - 1]} \right]. \end{aligned}$$

Under PT ($DS = 1$) we, therefore, obtain

$$\frac{\partial X}{\partial t} = \frac{[tf + s]^\beta [Y - X]^{\beta-1}}{\phi_1 D} \phi_{1t} [\beta w(p) - [Y - X] w'p'] v(\phi_1 - 1), \tag{A.8}$$

which takes the sign of $-\phi_{1t}$. Under RD ($DS = 0$) we obtain

$$\frac{\partial X}{\partial t} \begin{matrix} \geq \\ \leq \end{matrix} 0 \Leftrightarrow \phi_{1t} \begin{matrix} \leq \\ \geq \end{matrix} \frac{f [\gamma - \beta] \phi_1 [\phi_1 - 1]}{[tf + s] [\gamma \phi_1 - \beta [\phi_1 - 1]]} < 0.$$

□

References

Ali, M. M., Cecil, H. W., & Knoblett, J. A. (2001). The effects of tax rates and enforcement policies on taxpayer compliance: A study of self-employed taxpayers. *Atlantic Economic Journal*, 29(2), 186–202.

Alm, J., Sanchez, I., & De Juan, A. (1995). Economic and noneconomic factors in tax compliance. *Kyklos*, 48(1), 3–18.

al-Nowaihi, A., & Pyle, D. (2000). Income tax evasion: A theoretical analysis. In Z. MacDonald & D. Pyle (Eds.), *Illicit activity: The economics of crime, drugs, and tax fraud* (pp. 249–266). Aldershot: Ashgate.

al-Nowaihi, A., Bradley, I., & Dhami, S. (2008). A note on the utility function under prospect theory. *Economics Letters*, 99(2), 337–339.

Apesteguia, J., & Ballester, M. A. (2009). A theory of reference-dependent behavior. *Economic Theory*, 40(3), 427–455.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, 27(1), 173–196.

Bell, D. E., & Fishburn, P. C. (2000). Utility functions for wealth. *Journal of Risk and Uncertainty*, 20(1), 5–44.

Bernasconi, M., Corazzini, L., & Seri, R. (2014). Reference dependent preferences, hedonic adaptation and tax evasion: Does the tax burden matter? *Journal of Economic Psychology*, 40, 103–118.

Bernasconi, M., & Zanardi, A. (2004). Tax evasion, tax rates and reference dependence. *FinanzArchiv*, 60(3), 422–445.

- Bruhin, A., Fehr-Duda, H., & Epper, T. (2010). Risk and rationality: Uncovering heterogeneity in probability distortion. *Econometrica*, 78(4), 1375–1412.
- Camerer, C. (2000). Prospect theory in the wild: Evidence from the field. In D. Kahneman & A. Tversky (Eds.), *Choices, values, and frames* (pp. 288–300). Cambridge: Cambridge University Press.
- Clotfelter, C. T. (1983). Tax evasion and tax rates: An analysis of individual returns. *Review of Economics and Statistics*, 65(3), 363–373.
- Crane, S. E., & Nourzad, F. (1986). Inflation and tax evasion. *Review of Economics and Statistics*, 68(2), 217–223.
- Dell’Anno, R. (2009). Tax evasion, tax morale and policy maker’s effectiveness. *Journal of Socio-Economics*, 38(6), 988–997.
- Dhmi, S., & al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior and Organization*, 64(1), 171–192.
- Eide, E. (2001). Rank dependent expected utility models of tax evasion. Working Paper 27/2001, International Centre for Economic Research.
- Feinstein, J. S. (1991). An econometric analysis of income tax evasion and its detection. *RAND Journal of Economics*, 22(1), 14–35.
- Fisman, R., & Wei, S.-J. (2004). Tax rates and tax evasion: Evidence from “missing imports” in China. *Journal of Political Economy*, 112(2), 471–496.
- Friedland, N., Maital, S., & Rutenberg, A. (1978). A simulation study of income tax evasion. *Journal of Public Economics*, 10(1), 107–116.
- Gordon, J. P. F. (1989). Individual morality and reputation costs as deterrents to tax evasion. *European Economic Review*, 33(4), 797–805.
- Harrison, G. W., & Rutström, E. E. (2009). Expected utility theory and prospect theory: One wedding and a decent funeral. *Experimental Economics*, 12(2), 133–158.
- Hashimzade, N., Myles, G. D., & Tran-Nam, B. (2013). Applications of behavioural economics to tax evasion. *Journal of Economic Surveys*, 27(5), 941–977.
- Isoni, A. (2011). The willingness-to-accept/willingness-to-pay disparity in repeated markets: Loss aversion or ‘bad-deal’ aversion? *Theory and Decision*, 71(3), 409–430.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–292.
- Kamdar, N. (1995). Information reporting and tax compliance: An investigation using individual TCMP data. *Atlantic Economic Journal*, 23(4), 278–292.
- Kim, Y. (2003). Income distribution and equilibrium multiplicity in a stigma-based model of tax evasion. *Journal of Public Economics*, 87(7–8), 1591–1616.
- Kim, Y. (2005). Audit misperception, tax compliance, and optimal uncertainty. *Journal of Public Economic Theory*, 7(3), 521–541.
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica*, 79(3), 651–692.
- Köbberling, V., & Wakker, P. P. (2005). An index of loss aversion. *Journal of Economic Theory*, 122(1), 119–131.
- Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4), 1133–1165.
- Masatlioglu, R., & Raymond, C. (2016). A behavioral analysis of stochastic reference dependence. *American Economic Review*, 106(9), 2760–2782.
- Neilson, W., & Stowe, J. (2002). A further examination of cumulative prospect theory parameterizations. *Journal of Risk and Uncertainty*, 24(1), 31–46.
- Piolatto, A., & Trotin, G. (2016). Optimal income tax enforcement under prospect theory. *Journal of Public Economic Theory*, 18(1), 29–41.
- Pommerehne, W. W., & Weck-Hannemann, H. (1996). Tax rates, tax administration and income tax evasion in Switzerland. *Public Choice*, 88(1–2), 161–170.
- Poterba, J. (1987). Tax evasion and capital gains taxation. *American Economic Review*, 77(2), 234–239.
- Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3), 497–527.
- Rablen, M. D. (2010). Tax evasion and exchange equity: A reference-dependent approach. *Public Finance Review*, 38(3), 282–305.
- Rees-Jones, A. (2014). Loss aversion motivates tax sheltering: Evidence from U.S. tax returns. doi:10.2139/ssrn.2330980

- Reinganum, J., & Wilde, L. (1986). Equilibrium verification and reporting policies in a model of tax compliance. *International Economic Review*, 27(3), 739–760.
- Sadiraj, V. (2014). Probabilistic risk attitudes and local risk aversion: A paradox. *Theory and Decision*, 77(4), 443–454.
- Santos-Pinto, L., Bruhin, A., Mata, J., & Åstebro, T. (2015). Detecting heterogeneous risk attitudes with mixed gambles. *Theory and Decision*, 79(4), 573–600.
- Schmidt, U., & Zank, H. (2008). Risk aversion in cumulative prospect theory. *Management Science*, 54(1), 208–216.
- Trotin, G. (2012). Solving the Yitzhaki paradox: Income tax evasion and reference dependence under prospect theory. AMSE Working Paper No. 2012-38.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323.
- Yaniv, G. (1999). Tax compliance and advance tax payments: A prospect theory analysis. *National Tax Journal*, 52(4), 753–764.
- Yitzhaki, S. (1974). A note on income tax evasion: A theoretical analysis. *Journal of Public Economics*, 3(2), 201–202.