Can a ‘veil of ignorance’ reduce the impact of distortionary taxation on public good valuations?

Morgan Beeson 1 · Susan Chilton 1 · Michael Jones-Lee 1 · Hugh Metcalf 1 · Jytte Seested Nielsen 1

Published online: 21 June 2019 © The Author(s) 2019

Abstract
Monetary valuations for mortality and health risk reductions, elicited through stated preference surveys, have been shown to include values placed on others’ welfare. Less attention has been paid to how these altruistic concerns, and hence willingness to pay (WTP), are affected when the public good is to be provided through taxation. We set up a theoretical framework to investigate this issue and show that when we move away from a first-best tax, WTP becomes distorted. However, once a ‘Veil of Ignorance’ is introduced into the model, this distortionary impact is reduced. We test this empirically using a laboratory experiment and confirm our theoretical findings, generating WTP values that are closer to the social optimum.

Keywords Public good · Distortionary taxation · Veil of ignorance

JEL Classifications C91 · D61 · D63 · D64 · D82 · H23

1 Introduction

Monetary valuations for mortality and health risk reductions have been shown to include values placed on others’ welfare. Stated preference surveys have found public valuations of risk reductions to be either greater than (Viscusi et al. 1988; Arana and Leon 2002) or less than (Johannesson et al. 1996; de Blaieij et al. 2003; Hultkrantz et al. 2006; Andersson and Lindberg 2009; Svensson and Johannson 2010; Gyrd-Hansen et al. 2016) private valuations. One explanation for this variation is that preferences over others’ welfare can take a number of different forms. For example, Jones-Lee (1991, 1992) shows that optimal provision of safety expenditure is necessarily higher when altruism is purely safety-focussed and lower when altruism is purely wealth-focussed.

Morgan Beeson
Morgan.Beeson@newcastle.ac.uk

1 Newcastle University, Newcastle upon Tyne, UK
Less attention has been paid to the how these altruistic concerns, and hence WTP, are impacted when a tax is used as a payment vehicle for valuations. Many theoretical studies assume public goods are funded through first-best lump-sum taxes (Bergstrom 1982; Jones-Lee 1991, 1992; Johansson 1994). Empirically, Gyrd-Hansen et al. (2016) found that WTP for own and others’ safety is higher under first-best taxation than uniform taxation implying that the payment vehicle affects peoples’ valuation.

The aim of this paper is to further investigate the impact of different systems of taxation on willingness to pay (WTP) when respondents are purely altruistic.1 Extending Bergstrom (2006)’s framework, we present a benchmark model of WTP under first-best lump-sum taxation. This model infers that WTP is equal to private benefit for both pure self-interest and pure altruism. This result aligns with the theoretical result of Bergstrom (1982) who shows that optimal provision of safety expenditure is equivalent for pure self-interest and pure altruism if safety expenditure is funded through first-best lump-sum taxation. We then relax the assumption of first-best taxation and find that WTP values become distorted.2 Our model infers that the distortion of WTP values is driven by an interaction between purely altruistic preferences and the system of taxation.

With the aim of reducing the distortion of WTP, we introduce a novel mechanism in which respondents are placed behind a ‘veil of ignorance’ (VoI). Behind a VoI respondents are uncertain of their position in society and are therefore uncertain of their wealth, private benefit from the public good, and of how the system of taxation affects them personally. Respondents do know the distribution of wealth and private benefits, and the system of taxation in place. We find that in theory, the condition for maximising expected utility is equivalent for pure altruism and pure self-interest. This means that valuations made behind a VoI are no longer distorted by taxation.

We test our theoretical findings by observing the impact of different systems of taxation on valuations using an economic decision-making experiment. Our experiment extends the design in Messer et al. (2010) and elicits valuations for laboratory public goods both in front of and behind a VoI under different systems of taxation. In accordance with our theoretical model, we find empirically that valuations made in front of a VoI become distorted when taxation is no longer first-best. WTP is therefore dependent on the system of taxation. However, behind a VoI, we find no distortionary effect of taxation. This work has some important implications for two areas of empirical application: (i) stated preference surveys, and (ii) support of public good provision provided through the regulatory framework. We return to these points and discuss them in more detail in Section 5.

The rest of the paper is as follows: in Section 2, we set out our model expressing the relationship between pure altruism and distortionary taxes in determining WTP valuations. We then present our VoI mechanism as a method for reducing the impact of

---

1 We follow Jones-Lee (1991)’s definition of “pure altruism,” that if an individual has an altruistic concern for another then they respect that person’s preferences. For example, Person A’s marginal rate of substitution of Person B’s wealth for Person B’s safety is the same as Person B’s marginal rate of substitution of her own wealth for her own safety. This is equivalent to “benevolence” as described by Bergstrom (1982). For the sake of consistency, we shall use the term “pure altruism” for the remainder of this paper.

2 In the public finance literature distortionary taxation is often thought of as taxation which alters economic behaviour (Ihori 2017). For example, a tax on labour may impact upon the supply of labour and tax on consumption may impact upon consumption. In this paper, we consider distortionary taxation as distorting the value of a public good as respondents take into account the distributions of both the costs and benefits in their decision.
distortionary taxes. From this model we derive two hypotheses to be tested. In Section 3, we describe the experiment designed to test these hypotheses and to demonstrate that the new mechanism can generate theoretically consistent responses in an empirical setting. Section 4 presents the results. In Section 5 we discuss our findings and draw some conclusions.

2 Theory

2.1 First-best taxation

Following the framework of Bergstrom (2006), consider an economy of two individuals with standard preferences indexed \( i \) and \( j \). The individuals, who have incomes of \( y_i \) and \( y_j \), have the opportunity to fund a public good which provides separate private benefits on which the individuals place values of \( x_i \) and \( x_j \). The provision of the public good comes at an additional cost, requiring the individuals to pay \( c_i \) and \( c_j \), where \( c_i + c_j = C \).

Each individual values the other’s welfare and, as a pure altruist, respects the other’s preferences. Thus, purely altruistic utility functions, \( u_i \) and \( u_j \), reflect both own private utility and the private utility of the other individual. The private utility of the other individual is given a weight of \( \alpha_i \) and \( \alpha_j \) measuring the degree of altruism felt toward the other individual. Thus the purely altruistic utility function for Individual \( i \) is given by

\[
\begin{align*}
  u_i &= y_i + \alpha_i y_j + z(x_i - c_i + \alpha_i(x_j - c_j)), \\
  &\text{where } z = 1 \text{ if the public good is funded, and } 0 \text{ otherwise. This functional form allows for pure self-interest (}\alpha = 0\text{) and pure altruism (}\alpha > 0\text{).}
\end{align*}
\]

It is agreed that the decision of whether to fund the public good shall be left to the rules of benefit-cost analysis (BCA). That is, if the sum of the benefits exceeds the sum of the costs then the public good will be provided. As Eq. (1) demonstrates, each individual has correctly accounted for both the benefits and costs to the other individual so that the problem of Bergstrom (2006)’s ‘naive benefit-cost analyst’ is avoided and there is no issue of double-counting in the form Milgrom (1993) describes.

To complete the cost-benefit calculation, WTP for the public good is elicited from each individual. WTP for Individual \( i \) is the maximum cost, \( c_i \), such that

\[
\begin{align*}
  y_i + \alpha_i y_j &= y_i + \alpha_i y_j + x_i - c_i + \alpha_i(x_j - c_j), \\
  &\text{That is, her social utility with the public good (}z = 1\text{) equals her social utility without the public good (}z = 0\text{). Rearranging Eq. (2) gives}
\end{align*}
\]

\[
\begin{align*}
  c_i &= x_i + \alpha_i x_j - \alpha_i c_j, \\
  &\text{where } c_i \text{ is Individual } i \text{'s WTP. Given Individual } i \text{ knows that Individual } j \text{ also gives his WTP, } c_j, \text{ Individual } i \text{ knows that Individual } j \text{'s WTP response will be}
\end{align*}
\]

\[
\begin{align*}
  c_j &= x_j + \alpha_j x_i - \alpha_j c_i,
\end{align*}
\]
Substituting Eq. (4) into Eq. (3) gives

\[ c_i = x_i + \alpha_i x_j - \alpha_i \left( x_j + \alpha_j x_i - \alpha_j c_i \right), \]  

(5)

from which it follows that \( WTP_i = c_i = x_i \); Individual \( i \)'s private benefit. In this case, the sum of the WTP values is equal to the sum of the private benefits. If the total cost paid is equal to \( c_i + c_j \) then the net benefit is zero to both individuals without the need for compensation.

This result implies that in the case of first-best lump-sum taxation, WTP is equal to the private benefit regardless of the strength of a respondent’s purely altruistic preferences. Each individual is simply willing to pay up to their own private benefit because the choice of WTP is independent of both the other individual’s WTP and tax contribution. This aligns with the result shown in Bergstrom (1982) and Jones-Lee (1991) who show that the optimal level of provision of safety expenditure is equivalent for pure altruism and pure self-interest.

2.2 Non-first-best taxation

In most countries taxation is involuntary and based on a predefined structure. This structure uses individual characteristics, often income, to determine the amount an individual will pay. The amount each individual contributes to public good provision is therefore based on the cost of provision and the system of taxation. Given aggregate WTP determines the upper threshold cost of provision, each individual’s WTP impacts upon everyone else’s tax contributions. This means that WTP will become a function of the system of taxation when respondents value others’ consumption of private goods.

Our benchmark model found optimal WTP values when public goods were funded through first-best lump-sum taxation. Consider another model which better represents public good provision funded through pre-defined tax structures. Now each individual’s choice will affect the amount the other pays. WTP will now be impacted by pure altruism through an interaction with how the public good is funded—even when costs and benefits have both correctly been accounted for, avoiding the double counting problem as noted above.

Let us say both individuals, before calculating their WTP, have agreed that each will pay a proportion, \( t_i \) and \( t_j \), of the total cost where \( t_i + t_j = 1 \) to ensure the full cost, \( C \), is covered. The social utility function for Individual \( i \) is now expressed as

\[ u_i = y_i + \alpha_i y_j + z \left( x_i - t_i C + \alpha_i \left( x_j - (1-t_i) C \right) \right), \]  

(6)

Again to find WTP we must equate the utilities with \( z = 1 \) and without \( z = 0 \) the public good as in Eq. (2) so that

\[ y_i + \alpha_i y_j = y_i + \alpha_i y_j + x_i - t_i C + \alpha_i \left( x_j - (1-t_i) C \right). \]  

(7)

This accounts for Individual \( i \) knowing that if she chose a larger WTP then this will increase the maximum amount the group are willing to pay and, by doing so, increase the maximum potential cost that Individual \( j \) may have to pay regardless of his choice.
It follows from Eq. (7) that
\[
WTP_i = t_i C^* = t_i \frac{x_i + \alpha_i x_j}{t_i + \alpha_i (1-t_i)}, \tag{8}
\]
where \(t_i C^*\) is Individual \(i\)'s WTP. This implies Individual \(j\) would pay \((1 - t_i) C^*\), the remainder of the cost.

There are two special cases to consider that generate a WTP value equal to the private benefit: pure self-interest and a first-best tax share. Under the assumption of pure self-interest \((\alpha = 0)\), Equation (8) simplifies to \(WTP_i = x_i\), the private benefit. This is the standard result that WTP should equal the private benefit and therefore the way in which the public good is funded does not impact upon responses. Thus, in a purely self-interested society the sum of WTP values is always the sum of private values, independent of the tax system in place. However, when preferences are purely altruistic, the tax system will impact WTP, posing a problem when measuring values empirically.

Under a first-best tax share, each individual contributes according to their share of the total private values. The tax share is the ratio of an individual’s private value to the sum of the private values, such that
\[
t^*_i = \frac{x_i}{x_i + x_j}. \tag{9}
\]

With the first-best tax share, Equation (8) simplifies to \(x_i\) regardless of the size of the weight placed on the others’ welfare. The sum of WTP is then equal to the sum of the private benefits. If the cost is equal to the sum of WTP, then individual tax contributions equal individual private benefits and thus net benefits are zero. This follows the results shown in the benchmark lump-sum tax model and in Bergstrom (1982) and Jones-Lee (1991).

The presence of the \(t_i\) terms in Equation (8) implies that the share of tax each individual pays will influence WTP responses. Taking derivatives of Equation (8), the WTP value with respect to \(t_i\) and \(t_j\) gives
\[
\frac{d WTP_i}{dt_i} = \alpha_i \cdot \frac{x_i + \alpha_i x_j}{[\alpha_i + t_i(1-\alpha_i)]^2} \geq 0, \quad \text{and} \tag{10}
\]
\[
\frac{d WTP_i}{dt_j} = -\alpha_i \cdot \frac{x_i + \alpha_i x_j}{[1-t_j(1-\alpha_i)]^2} \leq 0. \tag{11}
\]

Given the first-best result and the signs of the derivatives, the model predicts that if Individual \(i\)'s tax share increases above (below) the first-best rate then her WTP will be greater (lower) than her private value. This means that both the magnitude and sign of \(t_i - t^*_i\) determines the size and direction of the distortion to WTP. This effect is driven by the tax shares \((t_i, t_j)\) and the relative sizes benefits of the public good which as demonstrated in Equation (9) drives \((t^*_i, t^*_j)\).

Whilst in our model we assume the \(\alpha\) term is exogenous, a non-zero \(\alpha\) term could be based on: wealth, relative wealth, relative benefits, some combination of the
aforementioned, or any number of other motivating factors. As the $\alpha$ term would likely vary among members of a society, this could result in either over- or under-provision of the public good depending on the relative strengths of altruism. This is a problem as either socially potentially Pareto improving proposals could be rejected, or costly proposals could ultimately be accepted.\footnote{The result from Bergstrom (2006) shows that a private values compensation test is a sufficient condition for a social potential Pareto improvement thus if the sum of social WTP values is less than the sum of private WTP values then it is possible a socially potentially Pareto improving proposal is rejected.}

Based on these results, pure altruism may cause variations in public good provision and not necessarily generate a social optimal in a society of pure altruists. It is possible to use private values only, however as pure altruism is not the only additional motivation beyond pure self-interest, the use of private values alone would violate the principle of consumer sovereignty. Ideally, other forms of altruism and other-regarding preferences would be accommodated, whilst avoiding the inefficiencies associated with pure altruism. In the next section, we adapt and extend the model to meet these criteria.

2.3 Behind a ‘veil of ignorance’

Our proposed mechanism for reducing the distortion of WTP is based on Harsanyi (1953)’s Utilitarian argument that ethical decisions should be based on the interests of individuals who are both stakeholders in society but who are also in a position of impartiality. Impartiality is generated through uncertainty over the individual’s position in the society. This concept of impartiality is more commonly referred to as a ‘Veil of Ignorance’ (VoI), as coined by Rawls (1972). We follow Harsanyi’s conception of a VoI as it takes the Utilitarian position that BCA and WTP studies are based on and, importantly, assumes citizens’ preferences remain unchanged when placed behind a VoI, whereas Rawls (1972)’s concept does not. Harsanyi (1955) suggested that the weighted sum of the expected utilities would be equivalent to the social expected utility, which when maximised would deliver a social optimum.

When respondents are placed behind a VoI, the distributions of wealth and private benefit from the public good and tax system in place are known to them but they do not know their own wealth, benefit and tax share. To maximise utility, respondents must consider themselves in each position in society with equal probability. In a society of two positions, a pure altruist will consider the possibility of being in each position and feel altruistically towards the other position. This means that each position is weighted by the respondent’s private preferences for the position and the altruistic concern from the other position. If the altruistic concern is equal for both positions, then they are given equal weight, as if the individual concerned is purely self-interested.

Thus, consider the same society of two positions denoted 1 and 2, but now there are two states of the world. The first has Individual $i$ in Position 1 and Individual $j$ in Position 2, and the second has Individual $i$ in the Position 2 and Individual $j$ in Position 1. Both states of the world occur with a probability of a half. Incomes and private benefits are now tied to position rather than individual and are represented by numbered subscripts.

We allow the $\alpha$ terms to vary to allow for differing strengths of altruism based on the state of the world. This allows for individuals to show more or less concern for the
other based on their position within the society. Now each individual has two terms for altruism $\alpha_1$ and $\alpha_2$, where the subscript denotes the position they are in. For example, $\alpha_1^i$ and $\alpha_2^i$ are the weights Individual $i$ places on Individual $j$’s welfare when Individual $i$ is in Positions 1 and 2 respectively.

Each individual now faces the same problem of equating utility with and without the public good provided similar to that expressed in Equation (7), however utility is now in expectation as the state of the world is not certain. Under the assumption of risk neutrality, Individual $i$’s problem can be expressed as:

$$\begin{align*}
y_1 + \alpha_1^i y_2 + \frac{\alpha_2^i y_1 + y_2}{2} &= \frac{y_1 + x_1 - t_1 C + \alpha_1^i (y_2 + x_2 - (1-t_1)C)}{2} \\
&+ \frac{y_2 + x_2 - (1-t_1)C + \alpha_2^i (y_1 + x_1 - t_1 C)}{2}.
\end{align*}$$

Solving for $C$ we get:

$$C^* = \frac{(1 + \alpha_2^i)x_1 + (1 + \alpha_1^i)x_2}{(1 + \alpha_1^i) + (\alpha_2^i - \alpha_1^i)t_1}$$

Where $C^*_i$ is the upper threshold total cost Individual $i$ would prefer the group to not surpass, or the Maximum Acceptable Cost (MAC). MAC behind a VoI then implies a separate WTP value for each position, generated by multiplying the individual’s MAC value by a given position’s tax share. In front of a VoI, dividing WTP by tax share generates the individual’s MAC which reflects the implied preferences over total spending as expressed in Equation (8) that $WTP_i = t_i C^*$. Thus, MAC can be applied both in front and behind a VoI. 4

Differentiating Equation (13) with respect to $t_1$ gives:

$$\frac{dC^*_i}{dt_1} = -(\alpha_2^i - \alpha_1^i) \frac{(1 + \alpha_2^i)x_1 + (1 + \alpha_1^i)x_2}{(1 + \alpha_1^i) + (\alpha_2^i - \alpha_1^i)t_1}^2$$

If preferences are purely altruistic, without care for distribution, so that $\alpha_1 = \alpha_2 \geq 0$, then Equation (13) simplifies to $x_1 + x_2$, the sum of the private values, and Equation (14) simplifies to 0. This implies that the valuation is stable regardless of the tax share. If $\alpha_1 \neq \alpha_2$ then the tax term does not drop out of Equation (13) and the sign size of the tax effect is directly impacted by the difference between the $\alpha$ terms, but not the absolute sizes of them, as seen in Equation (14). As $\alpha_1 \to \alpha_2$ the tax effect diminishes as $dC^*_i/dt_1 \to 0$. Because only the relative sizes of the $\alpha$ terms matter, the effect of distortionary taxation will be reduced to capture only the preference over distribution. 5

4 The concept of a MAC is similar to Bergstrom (2006)’s idea of the appropriate question to ask in BCA to avoid double-counting: “what is the largest tax increase you would be willing to accept for you and those like you”? (Bergstrom 2006, pg. 348).

5 Again, under first-best taxation the valuation shown in Equation (13) collapses back to the sum of the private values and the derivative with respect to tax share shown in Equation (14) is 0.
applying a VoI, our model infers that the distortionary impact of taxation and pure altruism on valuations have been accounted for, whilst distributional preferences remain.

2.4 Hypotheses

We test two predictions of our model using a laboratory experiment: [1] in front of a VoI, the first derivative of WTP with respect to tax share is positive when respondents are purely altruistic from Equation (10); and [2] behind a VoI, the first derivative of MAC with respect to tax share is zero when respondents are equally altruistic across positions from Equation (14).

The first prediction is tested with Hypothesis 1 (H1):

\[ H_0 : \frac{d \text{WTP}_i}{dt_i} \leq 0 \]

\[ H_a : \frac{d \text{WTP}_i}{dt_i} > 0 \]

The null hypothesis is that as tax share increases, WTP either decreases or remains the same. The alternate hypothesis is that as tax share increases, WTP increases.

The second prediction is tested with Hypothesis 2 (H2):

\[ H_0 : \frac{d C^*_i}{dt_1} = 0 \]

\[ H_a : \frac{d C^*_i}{dt_1} \neq 0 \]

The null hypothesis is that as tax share increases, MAC \( (C^*_i) \) remains the same. The alternate hypothesis is that as tax share increases, MAC either increases or decreases.

A laboratory experiment provides the best environment to test our hypotheses as variation in tax distortion can be generated whilst controlling for other potentially confounding factors. The next section describes our experiment.

3 Experimental design

To test the hypotheses we extend the Random Price Voting Mechanism (RPVM), presented in Messer et al. (2010) and Messer et al. (2013). The RPVM, an extension of the Becker-DeGroot-Marschak mechanism (Becker et al. 1964), is designed to elicit efficient WTP values in a public goods voting context. Within this mechanism, a subject’s weakly dominant strategy is to offer their true WTP. We extend the mechanism in three ways: [1] to allow for the cost of the public good to be shared asymmetrically; [2] by applying a VoI to respondents; and [3] by altering the RPVM procedure to elicit MACs. Further, we elicit preferences both in front and behind a VoI. Eliciting preferences in both situations facilitates comparisons between WTP values under the two settings. Preferences are elicited for four levels of a public good delivering different
private returns under four different tax systems: first-best tax, uniform tax, proportional wealth tax, and progressive wealth tax. Each level of the public good delivers different relative returns and therefore has differing first-best rates so that each tax system generates different levels of tax distortion for each level of the public good.

3.1 Experiment protocol and game

At the beginning of each experimental session subjects were informed that during the session they could earn tokens and that their earnings would be dependent on choices made by themselves and the choices of others during the session. Tokens, worth £0.05 each, were converted into cash at the end of the session. Instructions were read out loud to subjects who completed 4 practice rounds before completing 22 actual rounds.

At the start of each round, subjects were randomly placed into groups of 2 and assigned one of 2 positions: one with a smaller endowment of 100 Tokens denoted Position S, and the other with a larger endowment of 150 Tokens, denoted Position L. Groups of 2 provide the simplest design to unambiguously test the distortionary effect of taxation.

Each group had the opportunity to fund a public good which provided positive returns with certainty. Subjects were given information on endowments, private returns from the public good, and the share of the cost each Position would pay if the public good was funded. In front of a VoI subjects knew which position they were assigned to and behind a VoI they did not. Once the subject had considered this information, they entered into a computer the largest amount they were willing for the group to pay in total for the public good (i.e. their MAC). Subjects entered their MAC for 22 rounds. The weakly dominant strategy is for a subject to enter her true MAC as, in that case, a subject should always vote yes for a cost she is willing to pay and vote no otherwise.

Following standard experimental procedures, a round was randomly selected to be played out for real and a cost was drawn at random for this round. For the round selected to be played, each subject’s MAC was compared to the total cost. Subjects with MACs greater than, or equal to, the total cost were counted as voting for funding the public good. Subjects with MACs less than the total cost were counted as voting against funding the public good. If the group voted for the provision of the public good, both group members paid their allocated share of the randomly drawn total cost from their endowment and received their return. If the group voted against the provision of the public good, neither group member paid and both were left with their initial endowment. Tied votes were resolved by a random draw.

---

6 Sessions were conducted using z-Tree (Fischbacher 2007) at the Behavioural and Experimental Northeast Cluster (BENC) experimental laboratory at Newcastle University for which students were invited to participate using the laboratory’s recruitment system. Eleven sessions were run in total, each taking approximately 80 min. Nine sessions had 20 participants, with one session consisting of 8 subjects and one of 16 subjects due to no-shows for a total of 204 subjects.

7 The groups changed between rounds so that subjects were not with the same group members in consecutive rounds and remained anonymous to the other subjects. The order of rounds was randomised to control for order effects.

8 Extensions to our design could consider groups of larger numbers if experimenters wished to introduce asymmetries in endowments and payoffs.

9 See Appendix for Figures 1 and 2 showing an example of the decision screens.

10 This took place at the end of the session to avoid wealth effects between rounds.
Subjects (n = 204) then completed a questionnaire to collect demographic data and were paid at the end of the session. They were guaranteed a minimum of £5.00 and on average earned £12.07.

3.2 Treatments

Table 1 describes the private returns for 4 different levels of public good provision and the returns to each position under a first-best tax share. For each level of the public good, subjects offered MAC responses both in front of and behind a V oI from both Position S (smaller endowment) and Position L (larger endowment). For every level of public good provision, each subject was randomly assigned to one of four tax systems described in Table 2. The four different tax systems were: first-best tax, uniform tax, proportional endowment tax, and progressive endowment tax. All subjects valued a given level of public good in front and behind a V oI under the same tax system. This allows for a direct comparison within subject, whilst ensuring that differences in valuations between tax types were not an artefact of treatment assignment.

The first-best tax, was calculated as a ratio between an individual’s return and the group total return, as described in Equation (9), and therefore varied across public goods with different relative private benefits. For the uniform tax the cost was split equally among group members so that tax share, was 0.5. The proportional endowment tax for Position S, was calculated using:

\[ t_{SP}^{prop} = \frac{y_S}{y_S + y_L} = 0.4, \quad \text{(15)} \]

i.e. Position S’s endowment as a ratio of the group total endowment. The tax share for the progressive endowment tax Position S, was calculated using:

\[ t_{SP}^{prog} = \frac{y_S^2}{y_S^2 + y_L^2} \approx 0.3, \quad \text{(16)} \]

i.e. Position S’s endowment squared as a ratio of the sum of the squared endowments of the group, resulting in the most sensitivity to relative endowments.

The size of tax distortion depends on the system of taxation in place and the first-best tax share. The purpose of the experiment is to test the impact of distortionary taxation on WTP responses for public goods. To isolate the effect of the tax distortion, the parameters for the experiment were chosen to give variation in the public goods being valued and hold the relative endowments and systems of taxation constant. The size of tax distortions depends solely on the relative sizes of the private returns across the four levels of public good provision. This best reflects a real world context where a given set of respondents, with different income levels, are asked to value varying levels of a public good.

Asymmetries in the private returns replicate a number of scenarios where the returns to public good provision favour one Position over another. However, the direction of the tax distortion depends on the tax share in relation to the first-best tax share. For example,

\[ \text{The mean subject age was 22.2 and 45% of subjects were male respondents. 51% of subjects were home students, and 31% postgraduate students.} \]
Public Good (Level) 1 favours Position L with regard to the private return, but depending on the system of taxation, may favour Position L (in the case of a uniform tax) or Position S (in the case of progressive taxation). Our model infers that the impact of altruism should be dependent on the size of the tax distortion not simply the relative size of the returns. Creating this variation in relative returns and tax distortion allows for the experiment to test the robustness of the model prediction in the results.

### 3.3 Estimation strategy

Each of the 204 subjects valued each of the 4 levels of public good provision described in Table 1 for a randomly assigned tax structure from Table 2. This generated a sample consisting of 816 observations for Position S, Position L and behind a VoI. To test the effect of distortionary taxation on valuations in front and behind a VoI we consider four dependent variables: WTP from Positions S ($WT{P}_S$) and L ($WT{P}_L$), the sum of the individual’s WTP ($\sum WTP = WT{P}_S + WT{P}_L$), and MAC given from behind a VoI ($MAC_v$). WTP values are calculated by multiplying the appropriate MAC by the individual’s tax share. $WTP_S$ and $WTP_L$ are used to test H1 and $MAC_v$ is used to test H2. $\sum WTP$ allows for additional comparisons of aggregate WTP and MAC behind a VoI.

In our analyses, variation in WTP and MAC values are explained by three independent variables: the private returns to each group member and the size of the tax distortion for Position S. The private returns variables capture variation in the level of public good provision. Tax distortion is calculated as:

### Table 1 Public good levels

| Public good level | Private return to Position S, $x_S$ | Private return to Position L, $x_L$ | First-best tax shares, $(t_S, t_L)$¹ |
|-------------------|-------------------------------------|-------------------------------------|--------------------------------------|
| 1                 | 40                                  | 60                                  | (0.4, 0.6)                           |
| 2                 | 50                                  | 50                                  | (0.5, 0.5)                           |
| 3                 | 60                                  | 40                                  | (0.6, 0.4)                           |
| 4                 | 50                                  | 0                                   | (1.0, 0.0)                           |

¹ First-best tax was calculated as $t_S = x_S/(x_S + x_L)$, $t_L = 1 - t_S$

### Table 2 Tax systems

| Tax system          | Endowment Position S, $y_S$ | Endowment Position L, $y_L$ | Tax shares, $(t_S, t_L)$ |
|---------------------|-----------------------------|-----------------------------|--------------------------|
| T1: First-Best tax ⁰| 100                         | 150                         | –                        |
| T2: Uniform tax     | 100                         | 150                         | (0.5, 0.5)               |
| T3: Proportional tax| 100                         | 150                         | (0.4, 0.6)               |
| T4: Progressive tax | 100                         | 150                         | (0.3, 0.7)               |

⁰ First-best tax is dependent on the relative private returns

ⁱ Uniform tax was calculated as $t_S = 0.5$

⁲ Proportional tax was calculated as $t_S = y_S/(y_S + y_L)$

⁳ Progressive tax was calculated as $t_S = y_S^2/(y_S^2 + y_L^2)$
\[
\Delta t_{s,i} = t_{s,i} \frac{x_{s,i}}{x_{s,i} + x_{l,i}},
\]

(17)

i.e. the difference between the tax share contributed by Position S and the first-best tax share for Position S. Tax Distortion (\(\Delta t_{s,i}\)) can take a positive or negative value dependent on whether Position S is paying more or less than the first-best share. We use tax distortion rather than tax share as our model predicts it is the sign and size of tax distortion that determines the sign and size of the distortionary effect on WTP.

We use OLS regression with standard errors clustered at the subject level to estimate the following models\(^{12}\) for each of the dependent variables described above:

Model (i) : \(WTP_{Small} = \beta_0 + \beta_1 x_{s,i} + \beta_2 x_{l,i} + \beta_3 \Delta t_{s,i} + \varepsilon_i.\) (18)

Model (ii) : \(WTP_{Large} = \beta_0 + \beta_1 x_{s,i} + \beta_2 x_{l,i} + \beta_3 \Delta t_{s,i} + \varepsilon_i.\) (19)

Model (iii) : \(\sum WTP = \beta_0 + \beta_1 x_{s,i} + \beta_2 x_{l,i} + \beta_3 \Delta t_{s,i} + \varepsilon_i.\) (20)

Model (iv) : \(MAC_{Veil} = \beta_0 + \beta_1 x_{s,i} + \beta_2 x_{l,i} + \beta_3 \Delta t_{s,i} + \varepsilon_i.\) (21)

For each model, the underlying private value of the public good is captured by the return to Positions S and L, respectively resulting in positive significant \(\beta_1\) coefficient for Model (i) and \(\beta_2\) coefficient for Model (ii). WTP should not be dependent on the other individual’s return and instead the altruistic component of WTP should be dependent on the size of the tax distortion term. If subjects are altruistic we would expect an insignificant \(\beta_2\) term for Model (i), an insignificant \(\beta_1\) term for Model (ii), and significant \(\beta_3\) terms for both. The coefficients for Model (iii) will be the sum of the coefficients for Models (i) and (ii). For Model (iv) the \(\beta_1\) and \(\beta_2\) terms will capture the respective returns to each position and are therefore both expected to be significant.

To test H1, i.e. whether WTP increases as an individual’s own tax share increases, we set up the hypothesis based on model hypothesis generated from Equation (10) as follows:

\[
H_0 : \beta_3 = 0, \\
H_A : \beta_3 \neq 0.
\]

For Model (i), presented in Equation (18), with \(WTP_S\) as the dependent variable; a rejection of the null with a positive coefficient for \(\beta_3\) would indicate subjects are responding to tax distortion in the expected manner for a pure altruist by offering a

\(^{12}\) We did a general to specific setup to test for non-linear relationships. We incorporated squared and cubic terms for all variables and found them to be insignificant and therefore opted for a model of linear relationships.
larger WTP to increase the probability the public good is provided. For Model (ii), presented in Equation (19), with $WTP_L$ as the dependent variable; a rejection of the null with a negative coefficient for $\beta_3$ would indicate subjects are responding to tax distortion in the expected manner for a pure altruist by reducing their WTP to reduce the probability the public good is provided at a high cost to their partner. The net effect will be captured in the $\beta_3$ coefficient in Model (iii) and could take either a positive or negative sign depending on the relative sizes of the $\beta_3$ terms for Positions S and L.

To test H2, that is, whether $MAC_v$ remains unaffected as tax distortion increases, we set up the hypothesis based on model hypothesis generated from Equation (14) as follows:

$$H_0 : \beta_3 = 0,$$
$$H_A : \beta_3 \neq 0.$$ 

For Model (iv), presented in Equation (21), with $MAC_v$ as the dependent variable, a rejection of the null would imply that valuations behind a VoI are responsive to tax shares.

## 4 Results

Table 3 presents the regression results for Models (i) – (iv). As expected, the underlying private value of the public good is captured by the return to Position S and L, respectively resulting in positive significant coefficients on $x_S$ for Model (i) and on $x_L$ for Model (ii). WTP is not dependent on the other individual’s return as the $x_L$ term for Model (i) and on $x_S$ term for Model (ii) are both insignificant. The altruistic component of WTP is instead dependent on the size of the tax distortion term which has significant coefficients for both Models (i) and (ii). This implies that values are consistent with the model’s prediction as

### Table 3 OLS regression results for Models (i) – (iv) expressed in Equations (18)–(21)

| Explanatory variable | (i) WTP$_S$ | (ii) WTP$_L$ | (iii) $\sum$ WTP | (iv) MAC$_V$ |
|----------------------|-------------|-------------|-----------------|-------------|
| $x_S$                | 0.440***    | -0.009      | 0.431***        | 0.501***    |
| (0.070)              | (0.095)     | (0.120)     | (0.137)         |             |
| $x_L$                | -0.046      | 0.819***    | 0.773***        | 0.727***    |
| (0.046)              | (0.038)     | (0.060)     | (0.058)         |             |
| $\Delta t_S$        | 12.406***   | -29.729***  | -17.324**       | 10.656      |
| (4.869)              | (5.325)     | (7.867)     | (7.349)         |             |
| $\beta_0$           | 14.578***   | -0.601      | 13.977*         | 13.620*     |
| (4.764)              | (5.214)     | (7.386)     | (8.213)         |             |
| $R^2$                | 0.035       | 0.257       | 0.142           | 0.189       |
| n                    | 816         | 816         | 816             | 816         |

*OLS regressions are run with standard errors clustered at the subject level

The explanatory variables are $x_S$ (return to Position S), $x_L$ (return to Position L), $\Delta t_S = t_S - \frac{x_S}{x_S + x_L}$ (the measure of tax distortion expressed in Equation 16), and $\beta_0$ is the constant term

*, **, and *** denote significance at the 90%, 95%, and 99% level, two-tailed test
private value acts as an anchor point whilst the altruistic effect on WTP varies with tax distortion. The next two sections present the results of the two hypotheses described above.

4.1 Hypothesis 1: In front of a Vol

Hypothesis 1 tests the model prediction that $\frac{dWTP_i}{dt}>0$ and the inverse relationship that $\frac{dWTP_j}{dt}<0$ if respondents are purely altruistic. To test H1, we use the regression results for Model (i). We find that tax distortion has a positive significant impact on mean WTP. This means that WTP increases when the difference between the actual tax share paid by S and first-best tax share for S increases. Based on the positive (12.4) and significant coefficient we reject the null hypothesis H1 ($p = 0.01$) for Model (i). Likewise, we consider the regression results for Model (ii). Here we find that tax distortion has a negative ($-29.7$) significant impact on WTP. This means that we can also reject H1 ($p = 0.00$) for Model (ii).

Rejecting the null hypotheses implies that subjects are responding to changes in distortionary taxation as our model predicts for a pure altruist. In our experiment, the distortionary tax effect is stronger when subjects are in the larger endowment position, suggesting the impact of altruism is stronger from Position L than S. This results in a significant ($p = 0.03$) net effect, measured by Model (iii), having tax distortion effect of $-17.3$. Overall the value of public goods with a distortionary tax that favours redistribution towards Position S (L) have significantly higher (lower) total values than at the first-best rate.

4.2 Hypothesis 2: Behind a Vol

Hypothesis 2 tests the model predictions that $\frac{dC^*_i}{dt}=0$ if respondents are purely altruistic and risk neutral. To test H2 we use the regression results for Model (iv) and find an insignificant impact of tax distortion on $MAC_{\text{Vol}}$. This insignificant coefficient (10.7) implies that we cannot reject the null hypothesis for H2 ($p = 0.15$). Our result implies that behind a Vol valuations are not impacted significantly by distortionary taxation even though the results from in front of a Vol indicate that subjects were acting altruistically towards their partners. This result follows the prediction of our model for risk neutral pure altruists and implies that by applying a Vol, distortion of valuations through taxation and altruism has been reduced to the point of insignificance.

5 Discussion and conclusions

This paper further considers WTP in an altruistic society by relaxing the assumption of a first-best tax. Our theoretical model, supported by our empirical results, shows that when the assumption of a first-best tax is relaxed, WTP responses become distorted away from private values. In order to resolve this problem, we introduce a novel mechanism which places individuals behind a Vol. We find theoretically that values will return to the private level and, for the range of tax systems tested empirically, valuations are no longer distorted by taxation.
Our findings imply that the value individuals place on a public good (e.g. a tax funded public investment in the reduction of air pollution generating a mortality risk reduction) is dependent on both the benefits and costs of public good provision to themselves and to others. It is therefore important that any valuation process based on societal values correctly incorporates these preferences.

Whilst the model here assumes public goods are funded directly through taxation, they also apply where there is a governmental intervention that results in additional costs to members of society—for example, a proposed government regulation or subsidy could result in the reduction of emissions by private firms. The benefits generated by this scheme through the mortality and/or morbidity risk reductions would be felt widely in society whilst the costs would be levied on customers through increased prices.

We also provide some further insights into the impact of different payment vehicles on responses in stated preference surveys. Our results confirm that if WTP is elicited from a purely self-interested perspective, the resulting survey values would be indicative of an optimum provision of the public good from both the perspective of a self-interested and a purely altruistic individual, given a first-best tax was utilised. In the case of mortality risk reductions, values would include preferences for own safety and own wealth, but would ignore preferences over others’ safety and others’ wealth. However, our empirical findings, and many previous stated preference surveys valuing mortality risk reductions (see Viscusi et al. 1988; Johannesson et al. 1996; Gyrd-Hansen et al. 2016, etc.), show that respondents do not always act in pure self-interest. Jones-Lee (1991, 1992) shows that preferences over others’ welfare is not limited to pure altruism. It is well accepted that preferences over others’ welfare are valid economic preferences and therefore, based on the principle of consumer sovereignty, should be included in valuations. Eliciting private values alone will then also violate the principle of consumer sovereignty.

Alternatively, given one could elicit values under a first-best tax system and also allow respondents to value others’ safety and wealth as well as their own, the findings of our benchmark theoretical model infer that under both purely altruistic and purely self-interested preferences, WTP would equal the private values. By doing so the problem of the naïve benefit-cost analyst as described by Bergstrom (2006) is avoided. Values collected under the assumption of a first-best tax are theoretically consistent with the rules of a compensation test as compensation is hypothetical.

However, it would be extremely difficult, if not impossible, to suggest credibly that a first-best tax vehicle would be implemented following a survey. Unfortunately, if other than a first-best tax is deployed, our results show that if respondents have preferences over others’ welfare, then values generated will be distorted. Additionally, if a first-best tax vehicle is used in the survey but not administered in practice, then the level of public good provided may be inefficient. This problem could be solved in part by eliciting MAC values in front of a VoI. That is, eliciting the maximum increase in societal spending based on a distribution of costs defined by the tax system to provide the provision of the public good. The MAC value acts as the natural extension to Bergstrom (2006)’s suggested question of “what is the largest tax increase you would be willing to accept for you and those like you” for avoiding the naïve benefit-cost analyst problem. As described in the theory section, WTP and MAC values are inextricably linked, but MAC values could provide a better framework for the elicitation of values for a complete proposal. Additionally, eliciting values behind a VoI provides an opportunity to control for the distortion of pure
Our findings suggest that by introducing a VoI, the impact of pure altruism will be limited to distributional preferences alone.

Accordingly, when designing a stated preference survey to elicit values from the public, consideration should be given to the following: [1] The choice of a tax as a payment vehicle means that WTP will be affected by how both the benefits and costs are distributed; [2] WTP and MAC values can be used to consider both private and public values at the individual and societal levels, with the proviso that MAC will almost certainly require some further development and piloting in a field setting before it could be used in a large scale survey; [3] Since all tax systems—except for first-best—will distort WTP, a VoI could be incorporated to mitigate this problem and, if successful, might indicate that once behind a VoI choice of a particular tax system is less important than in front of a VoI; [4] Establishing how a VoI might impact on or interact with other biases that are known to affect mortality and or morbidity valuation studies.

Appendix

Fig. 1 Example decision screen for respondents in front of a Veil of Ignorance (VoI). Notes: To enter a MAC subjects moved a slider on the screen to the amount they wanted to select. The slider, which started in a random position, had a minimum of 0 and the maximum amount was set such that no group member could go bankrupt. As the subject moved the slider along, a table underneath showed the number of tokens each group member would contribute and their net payoff if the public good was funded at the selected amount. By doing so, subjects could see the cost the other group members would pay, allowing them to take this into account in their decision and thereby avoiding any double counting.
Fig. 2  Example decision screen for respondents behind a Veil of Ignorance (VoI)

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.

References

Andersson, H., & Lindberg, G. (2009). Benevolence and the value of road safety. Accident Analysis and Prevention, 41(2), 286–293.

Arana, J. E., & Leon, C. J. (2002). Willingness to pay for health risk reduction in the context of altruism. Health Economics, 11(7), 623–635.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. Behavioural Science, (July), 226–232.

Bergstrom, T. C. (1982). When is a man’s life worth more than his human capital? In M. W. Jones-Lee (Ed.), The value of life and safety: Proceedings of a conference held by the Geneva association. Amsterdam: North Holland.

Bergstrom, T. C. (2006). Benefit-cost in a benevolent society. American Economic Review, 96(1), 339–351.

de Blaetij, A., Florax, R. J. G. M., Rietveld, P., & Verhoef, E. (2003). The value of statistical life in road safety: A meta-analysis. Accident Analysis and Prevention, 35(6), 973–986.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. Experimental Economics, 10(2), 171–178.

Gyrd-Hansen, D., Kjær, T., & Nielsen, J. S. (2016). The value of mortality risk reductions. Pure altruism—a confounder? Journal of Health Economics, 49, 184–192.
Harsanyi, J. C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy, 61*(5), 434–435.

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy, 66*(4), 309–321.

Hultkrantz, L., Lindberg, G., & Andersson, C. (2006). The value of improved road safety. *Journal of Risk and Uncertainty, 32*(2), 151–170.

Ihori, T. (2017). *Principles of public finance*. Singapore: Springer.

Johannesson, M., Johansson, P. O., & O’Connor, R. M. (1996). The value of private safety versus the value of public safety. *Journal of Risk and Uncertainty, 13*(3), 263–275.

Johansson, P. O. (1994). Altruism and the value of statistical life – Empirical implications. *Journal of Health Economics, 13*(1), 111–118.

Jones-Lee, M. W. (1991). Altruism and the value of other people’s safety. *Journal of Risk and Uncertainty, 4*(2), 213–219.

Jones-Lee, M. W. (1992). Paternalistic altruism and the value of statistical life. *Economic Journal, 102*(410), 80–90.

Messer, K. D., Poe, G. L., Rondeau, D., Schulze, W. D., & Vossler, C. A. (2010). Social preferences and voting: An exploration using a novel preference revealing mechanism. *Journal of Public Economics, 94*(3–4), 308–317.

Messer, K. D., Poe, G. L., & Schulze, W. D. (2013). The value of private versus public risk and pure altruism: An experimental economics test. *Applied Economics, 45*(9), 1089–1097.

Milgrom, P. (1993). Is sympathy an economic value – Philosophy, economics, and the contingent valuation method. *Contingent Valuation, 220*, 417–441.

Rawls, J. (1972). *A theory of justice*. Oxford: Clarendon Press.

Svensson, M., & Johansson, M. V. (2010). Willingness to pay for private and public road safety in stated preference studies: Why the difference? *Accident Analysis and Prevention, 42*(4), 1205–1212.

Viscusi, W. K., Magat, W. A., & Forrest, A. (1988). Altruistic and private valuations of risk reduction. *Journal of Policy Analysis and Management, 7*(2), 227–245.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.