Alternative Payment Vehicles in Contingent Valuation: The Case of Genetically Modified Foods

Andreas Kontoleon and Mitsuyasu Yabe and Laura Darby

Department of Land Economy, University of Cambridge, UK

2005

Alternative Payment Vehicles in Contingent Valuation:
The Case of Genetically Modified Foods

Andreas Kontoleon[1], Mitsuyasu Yabe[2], and Laura Darby[1]

[1]Department of Land Economy, University of Cambridge, UK
[2]Department of Agricultural and Resource Economics, Kyushu University, Japan

Jan 2005

Abstract
In this paper, a secondary consumer food survey is used to explore the role of the payment vehicle in contingent valuation. More specifically, the paper investigates the household willingness-to-pay in the UK for a GM and non-GM labelling program under two alternative payment vehicles: 1) a standard product tax, under which consumers must trade-off some of their personal income for the labelling program; and 2) a taxation reallocation scheme, whereby consumers must trade-off some amount of their household's taxation money that is currently spent on other government-funded goods. Contrary to previous valuation research, the willingness-to-pay under each vehicle is not found to be statistically significantly different, suggesting that in the case study investigated here, the marginal values of private income and other public goods in the UK are approximately equal.

1 – INTRODUCTION

1.1 The Importance of the Payment Vehicle in Contingent Valuation

Contingent valuation (CV), a survey-based methodology for eliciting consumers' valuations of non-market resources, has been applied by researchers and policy-makers for several decades (Jacobsson & Dragun, 1996:90).[1] While there is still some controversy over its application (see Hanemann, 1994; Diamond & Hausman, 1994), a number of issues associated with its use have been tested and loosely resolved, for example, by inclusion in best-practice guidelines (Jacobsson & Dragun, 1996:93; Arrow et al, 1993).

One issue that would appear to have been given relatively limited attention, however, concerns the payment vehicle (PV), and in particular the form of its application. Indeed, while a number of feasible PVs have been analysed, including income, sales and property taxes, trust fund payments, entry fees and changes in utility bills (Garrod & Willis, 1999:132), one important real-life PV has until recently been ignored by CV practitioners. The PV in question concerns a reallocation of taxation funds, a payment mechanism which would appear to be a common source of policy financing given the political unpopularity of imposing new taxes or charges.

The potential impact of a tax reallocation vehicle on CV results may have important implications for policy-makers basing their decisions on valuation research. Indeed, if the willingness-to-pay (WTP) for a particular non-market resource differs under a tax reallocation vehicle as compared to a more traditional PV, it is important that policy-makers intending to finance a resource by way of tax reallocation, base their decisions on studies using the correct payment scenario. Alternatively, if there is little difference between WTP values under the different vehicles, traditional studies may be able to provide policy-makers with approximate resource valuations.

To date, only one paper by Bergstrom et al (2004) has attempted to examine the tax reallocation PV and compare it to more traditional payment mechanisms. This paper attempts to add to this emerging literature, and in particular, answer Bergstrom et al's (2004:14) call for further empirical investigation into this unique valuation and financing approach.

1.2 Aims and Objectives

In this paper the primary aim is to compare and evaluate WTP measures for a non-market good under two different payment scenarios. The first is under a product tax, a more typical PV, where consumers must trade-off their personal income for the non-market good. The second is under a tax reallocation scheme where consumers must trade off some amount of their household's taxation money that is currently spent on other government-funded goods. This comparison is achieved using a case study on

[1] For a recent and full exposition of the contingent valuation methodology see Bateman et al (2002).

the WTP for the separation and labelling of genetically modified (GM) and non-GM foods, an issue of some interest given recent European Union (EU) legislation on GM labelling and traceability.

Through the case study, and use of secondary consumer food survey data, this paper addresses two further objectives. The first of these is to determine the WTP of households in the United Kingdom (UK) for the separation and labelling of GM and non-GM foods, and the second is to examine the particular explanatory factors that contribute to this WTP. The exploration of these secondary objectives enables this paper to add to a growing literature that applies valuation techniques to issues of biotechnology.

The overall structure of this paper is as follows. In Section 2 the paper discusses PVs and their role in CV methodology, as well as surveying the existing empirical literature regarding PV effects. A conceptual model is then developed based on the work of Bergstrom et al (2004) which explores how product tax and tax reallocation PVs affect the WTP for non-market goods. In Section 3, the paper introduces the GM case study, and explains why this is a useful and policy-relevant scenario under which to test PV hypotheses. Information about methodology, data and variables is provided in Section 4, and an examination of the estimation results presented in Section 5. Policy considerations and conclusions are discussed in Section 6.

2 – THEORY AND LITERATURE

2.1 The Role and Impact of the Payment Vehicle in Contingent Valuation

Although CV is often criticised, it continues to be widely researched and utilized to measure non-market values (Champ et al, 2002:591). Given the methodology's popularity, it is therefore important to understand how its major elements can impact upon the elicitation of valid and reliable WTP values. The PV is one such element in the application of CV as it provides the context in which respondents are to pay their offered bids (Morrison et al, 2000:407; Cummings et al, 1986:31). Indeed, while early CV literature focused only on the direct role of the resource in question, by assuming all other elements of the hypothetical market to be neutral (Hoevenagel, 1990), it is now recognized that the valuation process itself will depend on all elements of the CV scenario, including the PV (Jacobsson & Dragun, 1996:86). That is, respondents in CV studies do not value levels of provision of a resource in abstract, but rather value a complete policy package under which one cannot separate the value of the resource from the procedures by which the resource is provided and payment is made (Mitchell & Carson, 1989:124; Cummings et al, 1986:209).

Different WTP estimates from different PVs are therefore not an unexpected result (Jacobsson & Dragun, 1996:89), and there is persistent empirical evidence that the form of the PV has a significant impact on CV outcomes. For example, in an early PV study, Rowe et al (1980) found that the WTP for landscape values was significantly greater under an income tax than an entrance fee, while more recently, Bateman et al (1995) found respondents were also WTP more for wetland protection under an income tax than a contribution to a charity or trust fund.

Given that PVs can be expected to alter CV results, PV bias generally refers not to specific variations in offered values and PVs (Cummings et al, 1986:209), but rather where the PV is either misperceived or is itself valued in a way not intended by the researcher (Mitchell & Carson, 1989). For example, if a PV is implausible or objectionable to a respondent, they may treat the elicitation question as hypothetical and modify their true bid, or refuse to answer the valuation question in protest (Morrison et al, 2000: 408; Jacobsson &Dragun, 1996:89).

In an attempt to avoid bias, it is recommended that a PV meet a number of key criteria. Bateman et al (2002:4.25,8.21) explain that an effective PV must be credible, relevant, acceptable and coercive, as well as being carefully pre-tested before a full CV survey is carried out. They suggest that the simplest guideline is to use the PV most closely representing that which is likely to be employed by the real-world decision. Garrod and Willis (1999:132) emphasize that the chosen vehicle must have a plausible connection with the resource it is being used to value, and also be perceived to be "fair" and "equitable" in its incidence and in relation to those deriving benefits from the proposed policy. For example, Greenley et al (1981) in a CV study on improving recreation amenities found that the WTP under a sales tax was four times as much as under an increase in residential sewerage fees. This was partly due to escalating sewerage fees, but primarily because respondents felt it was unfair that tourists would be able to benefit from the proposed improvements under the latter scenario without contributing to the expense. According to Jacobsson & Dragun (1996:89) and Morrison et al (2000:407), PVs should also be appropriate to the cultural and institutional arrangements particular to the location of the study, as well as being perceived as directly linked to the provision of the resource in question (Hanemann, 1995:89).

CV practitioners tend to classify PVs into two broad categories (Bergstrom et al, 2004:2). Those measured by direct income changes where the payment is fixed regardless of the amount of the resource consumed, such as tax levies and license fees; and those measured by commodity price changes where the payment can be adjusted by altering the amount of the resource consumed, such as entrance fees and utility bills. Bateman et al (2002:4.13) suggest that the use of the latter category can pose problems in CV, as respondents may state a positive WTP for a resource change, but simply adjust the quantity consumed so that their total expenditure remains the same.

An alternative classification of PVs is to group them by voluntary vehicles, a particular example of Bergstrom et al's second category, and coercive vehicles, such as taxes, prices and fees (Bateman et al, 2002:4.13). Each of these categories has different implications for strategic behaviour and the possibility of PV bias.

Voluntary contributions, for example, provide an opportunity to free-ride, and respondents have an incentive to overstate their WTP to secure provision of a resource, and then decide at a later point whether or not to provide the pledged funds (Carson, 1997:1503). Hanemann (1995:90) explains that most appeals for donations are not designed to elicit maximum WTP values and therefore represent a flawed paradigm for CV. This is because voluntary PV studies often simply ask for some donation, without identifying the particular consequences or outcomes that would not be able to occur without it. Garrod and Willis (1999:157) suggest respondents may actually be likely to contribute less than their true bid value when faced with a voluntary PV, as they may perceive it as an ineffectual mechanism, with only a small proportion of funds going to the resource in question, as opposed to management costs. This was seen to be a contributing factor in Jacobsson and Dragun's (1996) valuation of wildlife protection, where respondents were WTP significantly higher bids under an income tax than a donation to a trust fund. Hidano and Kato (2002) in their landscape amenity study explain that bid values might also be significantly lower under voluntary PVs because respondents feel it is inequitable if not everyone benefiting from a policy is compelled to pay. Bateman et al (2002:4.13) suggest that there is some consensus that voluntary PVs should generally be avoided, particularly due to the "insurmountable problems of free-riding." However, they suggest that ultimately the choice of the PV will depend on the resource being studied and the context in which it is likely to be provided.

Coercive PVs, while generally more incentive compatible than voluntary vehicles (Carson, 1997:1503) are still prone to PV bias. For example, taxes may be seen as more credible than donations, particularly if respondents believe that they ensure everyone pays and that funds are fully channeled. However, taxes might equally be seen as less credible and worthy only of a protest response if respondents focus on the exclusion of non-tax payers from the vehicle (Bateman et al, 2002:4.13), or do not trust that the funds raised will go to the resource in question (Green & Tunstall, 1999:247). Indeed, in the UK and some other countries, taxes are generally not hypothecated (Bateman et al, 2002:4.13, 8.21). This means a respondent may adjust her true WTP downwards after assessing the risk that a WTP for one good will simply be raised through a tax and then that sum will be applied to a different purpose (Green & Tunstall, 1999:247). Johnston et al's (1999) findings appear to support this hypothesis as they provide evidence that the WTP for a watershed management package is significantly greater under a tax constitutionally guaranteed to fund the package in question, than under a tax with no secure funding guarantee. This fits with Bateman et al's (1995) broader suggestion that as the expected probability of a resource being provided decreases, so too do WTP values.

2.2 A New Form of Payment Vehicle – The Tax Reallocation

Bergstrom et al (2004) have recently added a new form of PV to the CV literature, the tax reallocation (TR). The TR alters neither income nor commodity prices, but instead relies on a reallocation of the governments existing taxation funds. Bergstrom et al (2004:2) explain that this PV reallocates fixed expenditures to finance new non-market goods, thereby decreasing the amount of a household's taxation money that can be spent on other government-funded goods. The potential for a TR to be an appropriate PV in particular situations has been previously recognised in CV studies, even though it has not until recently been specifically tested. For example, in Morrison et al's (2000) study on the WTP for an environmentally friendly drainage pipe, 32% of the pre-test sample proposed that the government should pay for this project out of existing taxation funds.

Given that many non-market goods may in practice be financed by TRs rather than by imposing new and politically unpopular taxes, Bergstrom et al's (2004) study is highly policy relevant. This relevance is only increased given that their results suggest that welfare measures corresponding to a TR may not be influenced by nominal income, and therefore may have the potential to lead to socially desirable public good allocation decisions from a cost-benefit analysis standpoint (Bergstrom et al, 2004:12).

In their study on the TR mechanism, Bergstrom et al (2004) develop the theoretical welfare measures that correspond to this new PV, as well as examining the implications and limitations of valuing and financing resources using this technique. The conceptual model they develop will be fully explored in Section 2.3.

To test their model, Bergstrom et al (2004) use a case study of a groundwater protection program in the United States (US) to compare the dichotomous-choice WTP acceptance rate under a TR vehicle with the WTP acceptance rate under a more standard tax levy. The null hypothesis that the acceptance rates are the same is rejected, and the TR is found to generate a mean WTP value almost eighteen times higher than a tax levy[2]. Bergstrom et al (2004:12) explain that this provides evidence that respondents are more willing to trade-off other government-funded goods for the water quality program than they are to trade off their own disposable income.

2.3 The Conceptual Framework

The following framework is based on Bergstrom et al (2004:2-5), and briefly describes the theoretical structure of CV welfare estimation for a standard tax and for a TR. An intuitive interpretation of these welfare measures is also presented.

It should be noted that while the case study investigated in this paper involves a product tax to pay for GM food separation and labelling, this tax is placed on what was uncovered to be a completely inelastic commodity, supermarket food[3]. Therefore, although technically a household could adjust the amount of

[2] Similar results are also obtained under Bergstrom et al's (2004) open-ended questionnaire.

[3] Only 8 respondents in the sample claimed they would reduce the amount of food purchased if presented with a product tax. Only 4 respondents claimed they would reduce this amount by more than 5%, with 3 of these appearing to be

food consumed once presented with this tax, this was not found to be an issue, and the tax faced by each household can be regarded as fixed. Accordingly, it is assumed in this paper that the product tax can be analysed in exactly the same manner as Bergstrom et al's (2004) treatment of their standard tax levy scenario.

2.3.1 Welfare Measure for a Standard Tax

The indirect utility function and the expenditure function provide the theoretical structure for welfare estimation in CV, with the aim being to estimate the changes in one of these functions. This change, once isolated, measures the net increment in income that is equivalent to, or compensates for, an exogenous change in the quantity or quality of a resource (Haab & McConnell, 2002:6). Where the change in the quantity or quality of a resource is positive, and utility is held constant, compensating surplus, or WTP, measures the maximum amount of income a household will pay to receive this new benefit (Haab & McConnell, 2002:6)[4]. Compensating surplus or WTP in the case of a standard tax can be modelled as follows (Bergstrom et al, 2004:2-4):

Let $u = u(X,Q,Z)$ be a household's preference function, where u is a direct utility function positively related to X, Q and Z, and $X = X_1,\ldots,X_m$ is a vector of private market goods. Let Q be the public good under investigation, in this case a program of GM and non-GM food labelling and separation, to be simply referred to as the 'labelling program.' Z is a composite commodity of all other public goods, with the exception of the labelling program, with unit price and value equal to the tax charged to a household.

Assume that all households maximise utility subject to the household budget constraint, $M = PX + Z$, where M is nominal income, and P is a price vector of private goods. Note that a given household's disposable income, $M_d = M-Z$, is assumed to be spent entirely on purchasing market goods, PX.

This utility maximisation problem yields a set of conditional household demand functions for the market goods, $x_i^* (P,Q,Z,M_d)$, which can be substituted into the direct utility function to derive a conditional indirect utility function:

$$u = v(P,Q,Z,Md) \qquad (1).$$

Inverting function (1) for M_d yields the following conditional expenditure function, which defines the minimum expenditure on private goods required to produce utility level u, given *P, Q and Z:*

$$e^* = Md = e^*(P,Q,Z,u) \qquad (2).$$

erroneous answers, given that these respondents claim they would reduce the amount of food purchased by significantly more than the amount of the tax.

[4] There are three other forms of welfare measures that can be calculated by a CV study. For a full outline of welfare measures in CV see Bateman et al (2002).

To minimize total expenditures on both private and public goods subject to a utility level of u, the restricted expenditure function is required:

$$e = e(P,Q,Z,u) \quad (3).$$

This function defines the total expenditure on both private and public goods required to achieve utility, *u*, given *P, Q and Z*, and is related to the conditional expenditure function as follows:

$$e = e(P,Q,Z,u) = e^*(P,Q,Z,u) + Z \quad (4),$$

where *e* is nominal income and *e** is disposable income.

The change in the labelling program can now be introduced through a change in *Q*. Assume that without the program, perceived general food quality is Q^0 and with the program it will improve to Q^1 where $Q^1 > Q^0$, and the supply of other public goods is held constant at Z^0. If there is no change in nominal income, WTP for the labelling program is given by compensating surplus (CS), which can be calculated using the restricted and conditional expenditure functions:

$$CS = e(P,Q^0,Z^0,u^0) - e(P,Q^1,Z^0,u^0)$$
$$CS = [e^*(P,Q^0,Z^0,u^0) + Z^0] - [e^*(P,Q^1,Z^0,u^0) + Z^0]$$
$$CS = e^*(P,Q^0,Z^0,u^0) - e^*(P,Q^1,Z^0,u^0) \quad (5),$$

where, $u^0 = v(P,Q^0,Z^0,M_d^0)$ is the utility level at which perceived food quality is at level Q^0 before introduction of the labelling program, and $M_d^0 = M - Z^0$.

Bergstrom et al (2004:4) explain that in accordance with standard consumer demand theory, this magnitude, CS, for a change in *Q* equals the marginal rate of substitution between the labelling program and disposable income. That is $MRS_{Q,Md}$.

2.3.2 Welfare Measure for a Tax Reallocation

Instead of trading off household expenditure to obtain the improvement in the labelling program, the TR vehicle requires a household to trade-off *Z*, the composite commodity of other public goods. To model this trade-off for the increase of Q^0 to Q^1, Bergstrom et al (2004:3) introduce a new welfare measure called the compensating tax reallocation (CTR). Using the same notation as section 2.3.1 above, this measure can be calculated as follows:

$$CTR = e(P,Q^0,Z^0,u^0) - e(P,Q^1,Z^1,u^0) \quad (6).$$

Given that under a TR, M_d^0 is equivalent for each of the right-hand side terms of (6), the conditional expenditure functions of each scenario are also equal. That is:

$$e^*(P,Q^0,Z^0,u^0) = e^*(P,Q^1,Z^1,u^0) \quad (7).$$

Equation (6) can therefore be re-written using equation (4) as:

$$CTR = [e^*(P,Q^0,Z^0,u^0) + Z^0] - [e^*(P,Q^1,Z^1,u^0) + Z^1] \quad (8).$$

And, equation (8) can be reduced to:

$$CTR = Z^0 - Z^1 \quad (9).$$

If equation (7) is then solved for Z^1 by the following equation:

$$Z^1 = Z^{1*}(P,Q^0, Q^1,Z^0,u^0) \quad (10),$$

then CTR can also be expressed as:

$$CTR = Z^0 - Z^{1*}(P,Q^0, Q^1,Z^0,u^0) \quad (11).$$

As disposable income is held constant in the case of CTR, CTR is equal not to the marginal rate of substitution between the labelling program and income, but rather the marginal rate of substitution between the program and the composite commodity of all other public goods, $MRS_{Q,Z}$. If CTR is greater than CS, $MRS_{Q,Z} > MRS_{Q,Md}$, this therefore implies that a household prefers an increase in disposable income or private goods to an increase in all other public goods. If CTR is less than CS it correspondingly implies the reverse, and a household will prefer an increase in its access to publicly-provided goods and services.

Bergstrom et al (2004:5) explain that the two compensating welfare measures derived above are largely determined by the relative marginal values of a household's existing bundle of public and private goods. For example, if a household's existing bundle of public goods has a high marginal value relative to private goods, perhaps because few public goods are currently provided, then CTR is likely to be less than CS. The reverse will likely be true if a household lives in an area where there is already a large supply of valuable public goods. This means that for a given jurisdiction, it is largely impossible to a priori predict the relative magnitudes of CTR and CS.

The following case study will test for differences between CTR and CS for a GM and non-GM separation and labelling program in the UK.

3 – THE CASE STUDY

The case study used to test for differences between CTR and CS in this paper is a CV study conducted by CSERGE[5] on the WTP in the UK for the labelling of GM and non-GM food. Specifically, the survey asks households how much they would be WTP for a program that ensures all foods are separated, transported and inspected with their GM content clearly labelled.

This is an important and policy-relevant area of study given a large number of countries and the EU have adopted, or plan to adopt, mandatory labelling for GM food (Phillip & McNeill, 2000). In the EU, of which the UK is a part, the first labelling requirements for GM foods were introduced in 1997 (Carter & Gruere, 2003:1). These initial controls have been strengthened a number of times (Sheridan, 2001), most recently in April 2004, with a stringent traceability and labelling scheme.[6] This type of labelling scheme is costly, and Noussair et al (2002:3) suggest that it is important to ensure that the spending on such a program is economically efficient. One way the economic efficiency can be examined is by comparing the cost of the required identity preservation systems and labelling regulations with an aggregate estimate of households' WTP for the labelling of GM foods (Kaye-Blake et al, 2002:2).

A number of valuation studies have previously examined GM-related issues however almost all of these examine the WTP a price premium for non-GM over GM foods, rather than for labelling per se. This is an important distinction to make because research suggests that consumers may have strong desires for labelling even if they are not generally opposed to consuming GM foods (European Commission, 2000, Boccaletti & Moro, 2000). A recent study by Kaye-Blake et al (2002) is an exception to this trend and investigates using a linked logit model, the WTP of New Zealand consumers for information on the GM status of their food. Their findings indicate a significant aggregate WTP for labelling of NZ$285m per annum, that easily exceeds the estimated costs of the labelling requirements mandated by the Australia New Zealand Food Authority (Kaye-Blake et al, 2002:14).

While their model had a fairly weak goodness-of-fit and a low level of predictive accuracy, Kaye-Blake et al's results suggest that the demand for GM food labelling is affected by the same factors that affect the demand for non-GM food[7]. For example, attitudinal variables were found to be just as important as economic or demographic factors (Kaye-Blake, 2002:14), consistent with the findings of McCluskey et al (2003) and Kaneko & Chern (2003).

Specific factors which played a significant role in increasing the WTP for labelling in Kaye-Blake et al's study included being female, more highly educated, in paid employment, consuming organic food, having some knowledge of biotechnology, and exhibiting little belief in the potential benefits of genetic engineering. Those respondents less likely to pay for labelling were unemployed or retired, in

---

[5] Centre for Social and Economic Research on the Global Environment based at the University of East Anglia. See http://www.uea.ac.uk/env/cserge/.

[6] This scheme ensures GM products can be traced through the supply chain, and that any foods that contain GM are labelled if the GM material exceeds a 0.9% threshold (Food Standards Agency, 2003).

[7] For recent CV studies that examine the demand for non-GM food see McCluskey et al (2003) and Kaneko & Chern (2003). Note that the demand for non-GM foods has also been addressed using experimental methods (see for example Lusk et al, 2001 & Huffman et al, 2002), and choice modelling techniques (see Burton & Pearse, 2002). A number of other relevant papers can be located on the websites http://www.agbioforum.org, and http://www.economia.uniroma2.it/conferenze/icabr/.

agricultural occupations, price conscious, able to perceive benefits in GM technology and with strong negative or positive levels of trust in large food manufacturing companies (Kaye-Blake et al, 2002:12).

Age and income were not included as independent variables, despite the fact that age is mentioned as a significant and negatively-signed explanatory factor in the demand for non-GM food, and that economic theory suggests income and WTP for labelling are likely to be positively correlated (Kaye-Blake et al, 2002:3). It should be noted however, that the effect of income on the demand for non-GM foods is largely inconclusive in empirical studies (Gamble et al, 2000), and was omitted in Kaye-Blake et al's model due to a large number of non-responses and after some unsuccessful preliminary analysis (Kaye-Blake et al, 2002:7).

There are a number of other variables that have previously been suggested to have an impact on the WTP for non-GM food that are not considered by Kaye-Blake et al. These include variables that may have a negative impact on WTP such as confidence in government (Kaneko & Chern, 2003), and household size (Medenhall & Evenson, 2002), as well as those that may have a positive effect such as strong risk perceptions (Kaneko & Chern, 2003), food safety concerns (McCluskey et al, 2001; Grimsaud et al, 2003), ethical concerns, and the presence of children in the household (Kaneko & Chern, 2003). To extend Kaye-Blake et al's (2002) study, some of these omitted variables are examined in this paper, as discussed in Section 4.

Kaye-Blake et al (2002:15) conclude their study by suggesting that their investigation cannot fully answer the question of whether the New Zealand labelling program is economically efficient. This is because while their investigation examines GM food labelling in a general sense, it is likely that the costs of labelling and the WTP for labelling vary by product. Indeed, the Cabinet Office's Strategy Unit (2003: 41) explains in a recent report that public attitudes towards GM foods are complex. They suggest that these attitudes, which are likely to influence the WTP for GM food labelling, vary markedly between GM crops and other GM foods, and between different types of GM produce and GM traits. A number of CV studies confirm this suggestion and find evidence that the WTP a price premium for non-GM food can vary substantially by product[8].

WTP values can also vary significantly between different countries[9]. Bredahl (2000) suggests that these cross-country variations result because attitudes towards GM are deeply embedded in more general nature and technology-based attitudes held by each country's consumers. In the UK, consumers' attitudes towards GM foods are relatively negative (Cabinet Office Strategy Unit, 2003:41). For example, Moon and Balasubramanian (2001) find that UK consumers are significantly more opposed to

[8] For example, Chern & Kaneko (2003) find that the average WTP to avoid GM vegetable oil, cornflakes, salmon and GM-fed salmon are 41, 25, 88 and 62% respectively higher than base product prices. Other studies that find differences between WTP for different products include McCluskey et al (2001) and Grimsrud et al (2002).
[9] For example, see Moon & Balasubramanian (2001), Chern et al (2002).

GM and more WTP a premium for non-GM foods than similar consumers in the United States. Meanwhile, Gaskell & Allum (2003:59) provide evidence that UK residents are more pessimistic about biotechnology than the majority of other European citizens, although there is some evidence this pessimism, along with a strong opposition to GM foods, may be slowly declining.

In summary, it is expected that this paper will uncover a positive WTP amongst UK consumers for the labelling and separation of GM and non-GM foods. This WTP is likely to be explained by a range of demographic, socioeconomic and attitudinal factors. A product tax (PT) in the form of higher food prices would seem to be a credible and reasonable PV with which to assess this WTP, particularly as it is coercive and would appear to be directly linked to provision of the resource in question. This PV, however, may not be entirely fair and equitable, due to the fact that even those consumers with no demand for food labelling are forced to contribute towards the program. Kaye-Blake et al's (2002) study on the WTP for food labelling uses a PV of higher food prices, as do the CV studies that examine the WTP for non-GM food.

A TR may also be a feasible PV in this case study, particularly given that the government could decide at any point to fund the EU-mandated labelling program out of existing taxation revenues. A TR vehicle would appear to have largely similar characteristics to imposing higher food prices, in that it seems reasonable and coercive, but has some chance of creating PV bias through limited fairness to respondents not concerned by the consumption of GM foods. In addition, either PV could generate additional PV bias if respondents do not trust food manufacturers, or the government to use or reallocate the total identified funds as promised.

4 – DATA, VARIABLES AND ESTIMATION METHODOLOGY

4.1 Data

The source of data for the case study is a mail questionnaire that was administered to a cluster sample of 2000 UK households in 2001 following a series of consultations, focus groups and pilot studies.[10] The survey methodology followed best practice survey prescriptions[11], and each survey included an information pack on the potential benefits and risks of GM foods so all respondents would have a reasonably equal footing in their understanding of the issues at stake.

The survey uses a dichotomous-choice elicitation question[12] to ask whether or not given a specified bid value, households are willing to vote for a program that ensures the GM content is clearly labelled on all

---

[10] The valuation survey section can be found in Appendix A. The remainder can be found in Kontoleon et al (2002).

[11] For example, the survey closely adhered to the 'Total Design Method' of Dillman (2000) and recommendations of Mangiore (1999) to ensure sample representativeness and minimum non-response rates. See Kontoleon (2002) for further survey methodology and sampling details.

[12] See Hanley et al (1997:389) for a discussion of different elicitation methods and the advantages and disadvantages of the dichotomous choice technique. Note that the dichotomous choice technique is recommended by the NOAA panel on CV.

supermarket foods. As part of this question, each household has an equal chance of being presented with one of 8 bid values, £1, £3, £5, £10, £15, £20, £30, or £40 for every £100 spent on supermarket food; and one of two PV scenarios, either a PT or a TR. The latter split sample design provides data for measuring CS and CTR, and testing hypotheses about the effects of the two different PVs on WTP values.

Faced with the valuation scenario in the survey, a household can vote for or against the labelling program, or as recommended by the NOAA guidelines, can elect a "don't know" response (Haab & McConnell, 2002:21). The "no" responses are divided into two types – a respondent can indicate that they are not willing to pay the bid but may be willing to pay a lower amount for the program, or alternatively that they are against the program in an absolute sense and would not be WTP anything for its introduction. Those households who vote for the program are then asked how much food they would consume if the program were to go ahead, whilst those who vote against it in an absolute sense are prompted to explain the rationale behind their decision.

The survey also contains questions about households' food purchasing habits and environmental concerns. These questions were developed after focus groups and extensive reviews of GM literature, and are useful sections from which to draw attitudinal variables. Finally, the survey provides demographic information about households including respondent age, education, household size and income.

From the 2000 surveys, 581 were available in a useable form. This number reduced to 533 after using Microsoft Excel to account for a number of non-response problems. Households that provided no answer to the valuation question were removed from the sample, as were those who did not respond to the attitudinal questions, or the gender and age questions. Fifty respondents who did not provide income details were not removed from the sample given this high level of non-response. Instead they were allocated the average monthly sample income of £2906.90, after imputation efforts proved unsuccessful[13]. Of the remaining 533 responses, 285 are from the PT treatment, and 248 from the TR treatment. Issues of non-response bias and sample representativeness are considered in Section 5.1.

4.2 Variables

4.2.1 The Dependent Variables

For purposes of preliminary analysis and to examine the sensitivity of WTP values to alternative assumptions and methodologies in variable construction, four dependent variables (**WTP1 – WTP4**) were created in Excel from the survey data. Each of these is a binary variable, and takes the value of one if a household is WTP the bid value presented, and takes the value of zero otherwise. Construction of each variable treats “yes” responses identically, but differs with respect to the treatment of “no” and “don’t know” responses. A summary of the dependent variables is presented in Table 1.

TABLE 1: Dependent Variables

VARIABLE	DEFINITION
WTP1	Excludes “don’t know” responses. Takes the value of 1 for “yes” responses, and 0 for “no” responses.
WTP2	Excludes “don’t know” and absolute “no” responses. Takes the value of 1 for “yes” responses, and 0 for remaining “no” responses.
WTP3	Takes the value of 1 for “yes” responses, and 0 for “no” and “don’t know” responses.
WTP4	Excludes “don’t know,” indifferent and protest responses. Takes the value of 1 for “yes” responses, and 0 for remaining “no” responses.

The first dependent variable tested, **WTP1**, follows a common strategy of excluding “don’t know” responses to the valuation question (Garrod & Willis, 1999:188). As per Section 5.1, this would not appear to be too problematic in this case study, as the average characteristics of respondents excluded are not substantially different to those of the rest of the sample. **WTP1** then treats all “no’s” as zeros, regardless of the rationale that lies behind these responses.

WTP2 is similar to **WTP1** and excludes all “don’t know” responses, but differs to **WTP1** by also excluding those households that indicate they are absolutely against the labelling program. That is, those households who choose the second of the “no” response options. This construction technique follows the recommendations of Haab and McConnell (2002:133) and Pearce (2003:13) by removing respondents who appear not WTP any amount for the program either for reasons of indifference or protest.

WTP3 is constructed by allocating a zero value to all “no” and “don’t know” responses. This is in response to Carson et al’s (1998) and Groothuis and Whitehead’s (1998) suggestion that the majority of respondents who answer “don’t know” to CV surveys would select “no” if instructed to make a more definitive choice.

Finally, **WTP4** expands on the construction of **WTP2**. “Don’t know” responses are excluded as per **WTP2**, particularly given Wang’s (1997:220,224) claim that treating “don’t know” responses as “no’s” is unlikely to be a sound strategy in all situations and may result in significant underestimation of WTP values. However, rather than immediately excluding what appear to be indifferent and protest responses, the **WTP4** variable is constructed by more carefully examining the rationale behind these absolute “no” responses. In particular, households who indicate they are against the labelling program,

[13] An OLS model designed to predict income values for households with missing values was found to be a very weak fit with an overall explanatory R^2 of only 0.9%. This may be due to the lack of an occupational or employment variable in the model.

but explain that this is because the bid amount is too high, are included amongst the “no” respondents who initially indicate that they may support the program at some lower amount. Households that are against the program because they believe someone else should pay the cost, require more information to make a decision, or fail to provide a reason, are excluded from the sample and set aside as protest responses. Note that those households who believe someone else such as the GM food companies, non-GM food consumers, or in the case of the PT, the government, should pay for the labelling program, provide some evidence of PV bias. Finally, those households who indicate they are not worried about GM food, or believe that the funds should be spent on other public services, which may also indicate little concern for the impacts of GM, are treated as indifferent respondents. These households are excluded from econometric models, but are used to adjust WTP values, as per Haab and McConnell (2002:136) to accommodate for the fact that parts of the UK public may be indifferent towards the consumption and accordingly the identification of GM foods.

Initial estimation and analysis on the determinants of the WTP decision is conducted using a series of models with each of the four dependent variables. Given results are quite similar across the models in terms of magnitude and sign, this paper largely focuses on those models using **WTP4**. **WTP4** is chosen as its construction makes intuitive sense, and it makes the greatest use of the information contained in the survey responses. **WTP4** is also used to calculate WTP values in order to test the PV hypotheses to be found in Section 4.3, although these values are also calculated using the alternative WTP variables in order to examine sensitivity to variable construction.

4.2.2 The Independent Variables

The independent variables to be used in estimation are those demographic and attitudinal items included in the survey, which the analysis in Section 3 indicated might have a significant influence on the WTP decision. These variables are summarized in Table 2. Before examining these variables, however, the bid price variable deserves considerable attention. This variable, **BIDPR**, is the first and potentially most important variable included in the econometric models and takes the value of the particular bid presented to each household participating in the study. **BIDPR** is expected to have a significant and negative impact on the WTP decision. That is, as the bid value increases to higher amounts, households are less likely to be WTP for the labelling program under investigation. **BIDPR** is an important variable as it is included both in general models examining the determinants of WTP, as well as specific models that calculate WTP values for each PV sample.

<u>Demographic Variables</u>

Demographic variables examined in the WTP models include gender, age, education, income, number of children in the household and total household size. After preliminary analysis all demographic variables except age, income and household size, are constructed as dummy variables as per the definitions in Table 2. The age and income variables are constructed by allocating each respondent with the midpoint of their elected age or monthly income bracket, with missing income values replaced by the average sample income as outlined in Section 4.1.

TABLE 2: Independent Variables

VARIABLE	DEFINITION
BIDPR	Takes the value of the presented bid price - £1, £3, £5, £10, £15, £20, £30, or £40
GENDER	Takes a value of 1 for male respondents
AGE	The midpoint of the respondents age bracket
EDU	Takes a value of 1 if a respondent has a university degree or professional qualification
INCOME	The midpoint of the respondents income bracket
CHILD	Takes a value of 1 if the household contains any children aged 16 or under
HHOLD	The total number of household members
F1	Ethical Resistance – represents high scores in survey questions that indicate resistance to GM foods on ethical grounds
F2	Mistrust and Disbelief - represents high scores in survey questions that reflect a sense of mistrust with food standards and in the policies and practices of government, scientific and agricultural-industry bodies, as well as a lack of faith in the ability of humans to overcome any particular risks involved in the use of biotechnology
F3	Environmental Concerns - represents high scores in survey questions that indicate a strong pro-environment mind-set
F4	Cost and Bargain Concerns - represents high scores in survey questions that reflect a high level of price sensitivity
F5	Food Safety Concerns - represents high scores in survey questions that indicate high levels of concern with general food safety

In line with the results of previous studies, **GENDER, AGE, EDU** and **HHOLD** are expected to have a negative impact on the WTP decision, while the coefficient of **CHILD** is expected to be positive. Consistent with economic theory, **INCOME** is also expected to have a positive sign in the PT sample, although this has not consistently proved to be a significant variable in GM-related valuation studies. The **INCOME** variable in the tax reallocation sample has no strict a priori expectation, although if Bergstrom et al's (2004) study is indicative, the income variable is likely to be insignificant and negative. While Bergstrom et al (2004) do not explore the reasons behind this negative sign, it may make intuitive sense in this paper's case study. This is because poorer households pay less of the total tax revenues that would be used to fund the labelling program and may therefore be anxious for the reallocation mechanism to be used rather than having to face higher product prices that would impact upon them in a disproportionately large manner. This may make poor households more likely to agree to higher bid values under the TR scenario, especially as the TR version of the survey makes reference to the fact that under the alternative situation of higher prices, lower income households are particularly adversely affected.

Unfortunately, the survey data does not include occupation variables, which were found to be significant indicators of WTP in Kaye-Blake et al's (2002) study, or locational variables. The latter may be a useful determinant of WTP under the TR scenario as different areas are likely to display varying bundles of public goods, with the households in these areas consequently placing varying marginal values on increases in publicly-provided services.

Attitudinal Variables

A number of attitudinal and behavioural questions are available in the survey from which to construct additional explanatory variables. For the purposes of this paper, however, five previously constructed latent variables are used to analyse the impact of attitudes on the WTP for labelling decision. These variables were created using factor analysis[14] on 37 likert-scale questions in the survey, and are useful because they can effectively capture similarly-themed attitudes within the same variable, rather than including large numbers of separate variables, which might potentially lead to problems with multicollinearity or specification error. In addition, this methodology ensures reasonable accuracy given that the data on the raw attitudinal variables is somewhat unclear because it has been manipulated to complete the factor analysis.

The five latent variables quite comprehensively cover the range of attitudes that might impact upon the WTP for labelling decision, and are included with their descriptions in Table 2. Following previous findings **F1, F2** and **F5** might be expected to have a positive sign, whilst **F4** is expected to have a negative one. **F3** might reasonably be either positive or negative, depending on whether households see GM as risky or beneficial for the environment. The coefficient signs on the latent variables also might vary between the 2 different PV samples if the impact of attitudes on the WTP under a TR differs to that previously investigated under the standard PT or price rise.

4.3 Estimation Methodology

This paper follows the approach of Cameron and James (1987) and directly estimates WTP models[15]. These models are estimated for each of the PV scenarios using the probit technique[16]. All models are estimated in the LIMDEP econometrics package using the maximum likelihood method.

To analyse the determinants of whether or not households are WTP for the labelling program, the following binomial probit equation is used:

[14] See Kontoleon (2002) for a detailed description of the latent variable construction. Factor analysis is a statistical technique that is commonly used to isolate the factors underlying respondents' answers to a series of questions. In essence, factor analysis finds underlying commonalities in responses (Teisl et al, 2003:8). See Kontoleon (2002) and Nunes (2002) for more details.
[15] Another common alternative follows the approach of Hanemann. See Hanemann & Kanninen (1998).
[16] Note that the logit specification could have equally been used to produce highly comparable results (Gujarati, 1995:554).

$$P(Y = 1 \mid X) = \int_{-\infty}^{\sum \beta_k X_k} [\exp(-u^2/2)/\sqrt{2\pi}]du \quad = \Phi(\beta_k X_k)$$

Where Φ is the notation for the standard normal distribution, and X_k are the explanatory variables discussed in Section 4.2.

To calculate average household WTP, this paper estimates probit models containing solely the dependent variable, constant and bid price. Three such parametric models are estimated and discussed, each involving a different transformation of **BIDPR**. In order to examine the sensitivity of the WTP values to the parametric distribution assumptions, a fourth WTP value is also calculated as per Bateman et al (2002) and Haab and McConnell (2002), using the non-parametric Turnbull method. Once the WTP values are calculated, Thaye's two-tailed t-test (Larson, 1982) is used to test the following PV hypothesis at the 95% confidence level:

H_0: $WTP_{PT} = WTP_{TR}$; H_1: $WTP_{PT} \neq WTP_{TR}$,

Where WTP_{PT} is the mean WTP of a household under the PT scenario and WTP_{TR} is the mean WTP of a household when faced with a TR. If the null hypothesis is not rejected this provides evidence that WTP values are not statistically different under the two PVs. That is, CTR is equal to CS and a household equally prefers an increase in disposable income or private goods to an increase in all other public goods. If the null is rejected, it alternatively indicates that CTR and CS diverge, suggesting differences in the marginal values of public and private goods.

5 – ESTIMATION RESULTS

5.1 Summary Statistics

Before examining the estimation results from the probit models, it is informative to analyse the summary statistics for the sample of respondents. Table B1 of Appendix B presents the means and standard deviations of the household-specific variables, and also provides statistics by the two PV treatments.[17] Thaye's t-test is used to compare the means of each variable between these 2 treatments, and the null hypothesis that the means are equal cannot be rejected. It is important that the means of these demographic and attitudinal variables are similar to avoid complicating the PV hypothesis testing. Johansson (1999:14) explains that if variables are significantly different across the sub-samples, it is difficult to tell whether a hypothesis test result is reliable, rather than due to some systematic difference in household characteristics.

In order to assess the potential bias from omitting certain categories of respondents, average sample characteristics of households included in the construction of each WTP variable are compared against households who are excluded. These are also compared, where possible, against characteristics of respondents omitted from the sample of 533 households. The results for **WTP4** are contained in Table B2. It can be seen that the mean characteristics across the first two groups of households are comparable, suggesting that the **WTP4** sample is largely representative of the total useable sample under investigation. The characteristics of the sample originally excluded however, vary more widely, suggesting some level of non-response bias. This level of bias is difficult to determine conclusively, given that the means in this third sample were calculated using limited data with a number of missing values. Weighting procedures could be used to adjust the retained sample if it is believed to be unrepresentative of the whole (Bateman et al, 2002:5.38), however this approach is not taken here given the unavailability of information on true full sample or population characteristics.

5.2 Results of the WTP Model

LIMDEP is used to estimate initial runs of the probit models[18] using the four dependent variables in each PV scenario, with the independent variables described in Table 2. Both on goodness-of-fit grounds and the grounds of individual parameter effects, results across the models are reasonably comparable. For example, pseudo R^2 values across the four PT models vary between 18.52% and 20.34%, while those for the TR models are between 10.79% and 12.54%. The **BIDPR** variable has a significant and negative impact on WTP in all models, while the remaining explanatory variables are consistently signed across the models for each PV scenario.

For reasons discussed in Section 4.2.1, and on the basis of slight statistical superiority, this paper focuses largely on the results of the **WTP4** models. Table B3 of Appendix B presents the initial results of the **WTP4** models, while Table 3 presents the results following exploration and testing of some alternative specifications. In particular, the models in Table 3 include an age-squared term, **AGESQ**, to investigate how the WTP decision varies with the rate of change in age, as well as the natural log of the income variable, **LINCOME**. On suggestion of Bateman et al (2002:5.18) all variables are retained in the models. This is regardless of their significance and allows for a full comparison and discussion of the covariate effects under each PV scenario.

TABLE 3: Estimation Results

VARIABLE	ST Marginal Effect	P-value	TR Marginal Effect	P-value
Constant	0.1827	0.7313	0.5024	0.3299
BIDPR	-0.0211	0.0000**	-0.0100	0.0057**
GENDER	-0.0331	0.6861	0.0356	0.4050
AGE	-0.0315	0.0465**	-0.0300	0.0692*
AGESQ	0.0004	0.0136**	0.0003	0.0620*
EDU	-0.0729	0.3789	-0.0710	0.4145

[17] Note that for ease of interpretation, the means of CHILD and EDU are given using continuous variables rather than the dummies used in estimation.
[18] See Appendix C for an excerpt of the final LIMDEP command file

LINCOME	0.0840	0.0880*	0.0279	0.5711
CHILD	0.0400	0.7134	0.0012	0.9921
HHOLD	-0.0043	0.9060	0.0335	0.3984
F1	0.0885	0.0319**	0.1797	0.0000**
F2	0.1191	0.0112**	0.0776	0.0936*
F3	0.0294	0.5348	0.0868	0.0755*
F4	-0.0695	0.1717	0.0622	0.2375
F5	-0.0875	0.0904*	0.0361	0.5030
Pseudo R^2 Test significance $\sim X^2$ Prob > X^2 Prediction Success	22.35% 62.7678 0.0000 73.40%		13.72% 33.6505 0.0014 66.10%	

** Significant at the 5% level; * significant at the 10% level

5.2.1 Testing the WTP Model

Before discussing the results of the models in Table 3, it is important to examine the overall goodness-of-fit of the models[19]. In assessing this fit there are several summary statistics for the probit specification that can be examined. These are set out in Table 3 and suggest that the PT model has a somewhat superior fit to the TR model. The first statistic is McFadden's pseudo $R^{2,}$ which is calculated as 1-(L_{max}/L_0), where L_{max} is the value of the unrestricted loglikelihood function, and L_0 is the value of the function when all slope coefficients are set equal to zero. Considerable care needs to be taken when interpreting this statistic which has values of 22.35% for the PT model and 13.72% for the TR model. This is because there is no commonly accepted value for pseudo R^2 that denotes a satisfactory or well-defined model (Hanemann & Kanninen, 1998:344); although higher values are preferred and any value in single figures suggests WTP values show little in the way of distinguishable patterns (Bateman et al, 2002:5.18).

An alternative and preferable measure of fit is the X^2 test statistic also found in Table 3. This statistic, which corresponds to the F-statistic in linear regression, by testing whether all slope coefficients are equal to 0, has calculated and p values of 62.77 and 0.0000, and 33.65 and 0.0014, for the PT and TR models respectively, suggesting that both models as a whole are statistically significant.

From a summary point of view, it is also worthwhile to report the predictive success of the models, where a prediction, P, is correct when $P_j > 0.5$ and $Y_j = 1$, or $P_j \leq 0.5$ and $Y_j = 0$, and where Y_j is the observed WTP decision equal to one when an individual chooses to vote for the labelling program (Fomby et al, 1988:352). A classification table indicating the predicted and observed values of **WTP4**,

[19] Note that given the scope of this paper, and the complexity of the testing and corrective procedures required, it is assumed that the basic assumptions of the probit model hold. Details on probit model assumptions, and testing and correcting for these assumptions can be found in Aldrich and Nelson (1984), Maddala (1983) and Greene (1993).

is presented in Appendix C. Results suggest that both models are fair predictors of the WTP decision with 73.40% and 66.10% of cases in the PT and TR samples correctly classified.

5.2.2 Discussion of WTP Model Results

The estimation results in Table 3 are generally consistent with previous empirical findings and/or make intuitive sense, suggesting that the internal validity of the study is sound. In particular, the **BIDPR** variable is negative and significant in both models and the estimated coefficients for the remaining variables while not always significant have with only limited exception, the expected signs.

The **AGE** and **AGESQ** coefficients are negative and positive respectively, suggesting older respondents are less likely to be WTP for GM labelling at an increasing rate. These variables are significant at the 5% level in the PT model and at 10% for the case of the TR. In both models **EDU** is negatively, but insignificantly related to the WTP decision, while having a child has a positive impact. As expected, the coefficient on household size has a negative sign in the PT model, although this variable is also insignificant. The **HH** variable is insignificant and positive in the TR model, perhaps because the alternative option of paying higher food prices is more threatening to households with larger numbers of members and thereby higher levels of food expenditure. Contrary to expectations, **GENDER** has a positive coefficient in the PT model, but as per the TR model, this variable is highly insignificant.

Finally, in line with economic theory, **LINCOME** has a positive and significant impact on the WTP for labelling decision in the case of a PT, but is insignificant in the TR scenario. The coefficient of **LINCOME** in the TR model is positive in contrast to Bergstrom et al's (2004) findings, but given its small magnitude and high level of insignificance, this paper's results support their suggestion that welfare effects measured by a TR are not influenced by nominal income. As Bergstrom et al (2004:12) explain, this may be a desirable result from a social justice perspective, and while this is not true of this case study, may be particularly useful in evaluating policies that provide benefits to richer households while imposing costs on the poor.

The models presented in this paper also suggest that attitudinal variables are important in determining the WTP for labelling decision, with three significant variables in each of the two models. In the PT model **F1** and **F2** are positive and significant, indicating as expected, that households with higher ethical resistance to GM and mistrust in the institutions involved in biotechnology are more likely to be WTP for GM labelling. **F3** and **F4** are insignificant but consistent with expectations, while **F5** is surprisingly negative and significant. This latter result suggests that high levels of food safety concerns decrease the likelihood of a household being WTP for labelling. It is unclear why this might be the case, although perhaps food safety-conscious households may for example, see the potential for GM to improve nutritional value as outweighing any possible adverse health risks, and as a result are not WTP an additional premium for labelling.

In the TR model, **F1** and **F2** are positive and significant as per the PT model, as is **F3** which indicates strong environmental concerns. As expected **F5** has a positive impact on the WTP decision, however it is highly insignificant. Finally **F4**, while insignificant is also positive, suggesting that households with strong cost and bargain concerns are more likely to be WTP for labelling under a TR scenario. This may reflect these households being more likely to agree to their allocated bid values under the TR model, because this involves reallocation of their existing taxation payments rather than having to pay additional funds as per the alternative scenario.

It should be noted this paper is somewhat limited to using the attitudinal variables as defined by a previous survey analysis. To more fully investigate the impact of different attitudes and behaviours on the WTP decision, it is recommended that further investigation with different types and forms of these variables take place.

5.3 Calculation of WTP and Hypothesis Testing

5.3.1 Calculation of Average Household WTP

This section of the paper follows Vaughan et al (1999) and Crooker and Herriges (2004), in calculating a range of central tendency estimates for household WTP. Russell (2001:179) explains that with a simple dichotomous-choice survey, it is possible to choose among approximately twelve methods for inferring WTP. Each of these methods generally produces a different measure of central tendency which can differ by several magnitudes to other measures, with often no particular measure so obviously superior that the others can be rejected (Russell, 2001:179, Vaughan et al, 1999:7). In this paper three common parametric estimates are compared and contrasted in order to examine the sensitivity of these estimates to the underlying distribution of preferences and the estimation procedures employed[20]. In addition, a non-parametric model is employed to produce a mean WTP estimate that is robust against potential distributional mis-specification (Kristrom, 1990:138).

The first parametric model investigated is the linear probit model. Following Cameron and James (1987) and Hanemann and Kanninen (1998), a probit regression containing a constant and the **BIDPR** variable is implemented and the estimates recovered. The coefficients on the constant and **BIDPR**, α and β, represent point estimates of μ/σ and $-1/\sigma$ respectively, and allow the mean WTP, μ, to be quickly recovered. Given the symmetry property of the normal distribution, μ is also equal to the median household WTP (Buckland et al, 1999:110). The first column of Table 4 contains estimates of

[20] Note that Bateman et al (2002:5.19) explain that the best-fitting model when covariates are included does not have to make the same distributional assumptions as those used in a parametric model used to evaluate mean and median WTP.

mean/median WTP under the linear probit model for the **WTP4** models, as well as the models using the three alternative dependent variables[21].

TABLE 4: Mean and Median WTP Values

Variable	Linear Mean / Median	Log (BIDPR) Mean	Log (BIDPR) Median	Bound Median	Turnbull Mean
WTP4 – PT	£11.98	£116.56	£7.21	£9.06	£11.74
WTP4 – TR	£11.91	£8660.52	£7.44	£10.90	£11.27
WTP1 – PT	£5.59	£88.16	£3.77	£3.42	£10.18
WTP1 – TR	-£3.69	£92,394.12	£2.15	£1.07	£9.61
WTP2 – PT	£12.77	£111.80	£7.97	£10.31	£12.57
WTP2 – TR	£12.22	£11,290.78	£7.60	£11.35	£11.49
WTP3 – PT	£2.69	£85.60	£2.76	£2.68	£8.80
WTP3 – TR	-£8.23	£11,661.73	£1.59	£0.81	£8.43

Note that while the calculated values under **WTP4** and **WTP2** are similar both between the two PV scenarios and to each other, the values under **WTP1** and **WTP3** are quite different, and in particular, the values under the TR scenarios are negative. Indeed, negative WTP results can occur in linear models despite no negative bids being offered. Haab and McConnell (2002:94) explain that this is because the symmetry of the normal distribution places an equal mass of the distribution above and below the mean. In the case of the TR models utilising **WTP1** and **WTP3**, there are a large number of "no" responses to the low bids in the sample data[22], which causes the linear model to put a large mass in the negative WTP region and produce negative estimates of expected WTP. According to Haab and McConnell (2002:85,96) in situations where it is unreasonable for mean WTP to be negative, the probable cause of negative WTP values is functional form, or the bid price structure. That is, the set of bid values may fail to elicit responses that accurately trace out the probability that an individual will pay a given amount[23]. In this case study, a negative WTP value in the TR scenario would appear to be unreasonable, because it suggests that households would like to reduce the current share of the tax budget that goes to a GM labelling program. However, as this program is not currently funded by taxation revenues, this particular share of the tax budget is non-existent and therefore cannot be reduced.

To ensure average household WTP is a non-negative value for all models, the second of the parametric methodologies is used. This technique assumes that the relationship between WTP and the bid price is log-linear, and estimates a simple regression model, replacing **BIDPR** with **LBIDPR**, the natural log of the **BIDPR** variable. In essence, this technique transforms the bid variable so that negative values of

[21] Following model estimation in LIMDEP, all WTP values were calculated using Excel.
[22] In WTP1 and WTP3 52% and 55% of respondents respectively, respond "no" to bids equal to or less than £5. This is in comparison to a figure of 41% for WTP2 and WTP4.
[23] There are a number of ways to remedy this problem in the questionnaire design phase. See Haab and McConnell (2002) for further details.

WTP are impossible[24] (Buckland et al, 1999:111). Under this methodology the recovered estimates of μ/σ and –1/σ are used to calculate mean and median WTP using the following formulae (Haab and McConnell, 2002):

$$\text{Mean WTP} = e^{\mu+\frac{1}{2}\sigma^2}; \quad \text{Median WTP} = e^{\mu}$$

The calculated WTP results of this technique can be seen in columns 2 and 3 of Table 4. Note that the mean WTP values are significantly larger than those under the linear models, particularly in the case of the TR scenario. This reflects the "fat tails" problem, whereby unrealistically large estimates of mean WTP result from models that bound WTP from below (Haab & McConnell, 2002:92). Buckland et al (1999:111) explain that while the log transformation might adequately resolve the difficulty of fitting the lower tail of the WTP curve, it can create an even greater difficulty in the upper tail. That is, by restricting WTP to be positive, this technique forces too large a proportion of mass in the upper tail of the WTP distribution and can give rise to absurdly high estimates of WTP (Buckland et al, 1999:112, Haab & McConnell, 2002:95). This is particularly true of the TR models where there are a large number of "yes" responses to the highest bid value[25], resulting in the log-linear model assigning large amounts of weight to the WTP values in the upper tail of the distribution.

One way to avoid the "fat tails" problem is to use the median WTP value of the log-linear model. The median is a more robust measure of central tendency as it is largely unaffected by high bids in the upper tail of the distribution (Garrod & Willis, 1999:139, Jacobsson & Dragun, 1996:105). Furthermore, according to Garrod and Willis (1999:139) it is a useful measure because it indicates the amount of money which a one-person–one-vote system would allocate to the policy in question. A disadvantage of the median however, is that it is unsuitable for aggregating WTP values across populations (Mitchell &Carson, 1989:197), and so cannot be used to estimate the total WTP for GM labelling across the UK.

An alternative remedy to the "fat tails" problem is to identify a transformation variable that removes the probability of the range of WTP in the lower tail, but does not significantly alter the behaviour of the upper tail[26] (Buckland et al, 1999:112). For example, Haab and McConnell (1997:17) suggest that in most cases WTP is bounded below by zero and above by income, and recommend that these bounds be incorporated into estimating and calculating central tendencies of WTP. This paper follows Haab and McConnell's approach by creating a new bounded bid variable, **BOUND** equal to $\ln[\frac{Y_i - BIDPR_i}{BIDPR_i}]$,

---

[24] An alternative technique is to estimate the linear model and then truncate the curve at 0 when calculating WTP. This method, however, is not recommended as it is seen to be theoretically inconsistent and arbitrary. See Buckland et al (1999:111) and Haab and McConnell (1996:15).

[25] The % of households responding "yes" to the highest bid of £40 are 11.5, 15.8, 10.3 and 8.7% in the PT models for WTP1, WTP2, WTP3 and WTP4 respectively, compared to very high values of 35.3, 42.9, 31.6 and 42.9% in the TR models. Ideally these %s at the highest bid value should be close to 0.

[26] Rather than identifying a transformation variable, truncation can also be used following estimation to truncate the upper portion of the WTP curve. However, similarly to truncating at 0, this is not the recommended approach.

where Y_i is the i^th household's calculated annual income, and $BIDPR_i$ is the bid value offered to household i[27]. By estimating this model and retaining the coefficients μ/σ and $-1/\sigma$, on the constant and **BOUND** variables respectively, median household WTP can be calculated using the following equation (Haab & McConnell, 2002):

$$\text{Median WTP} = \frac{\overline{Y}}{(1+e^{\mu})}$$

Where $\overline{Y}$ is equal to average sample annual income. The values of the median household WTP using the bounded model are found in column 4 of Table 4, and are similar to those from the linear and log(**BIDPR**) median models, with the exception of the TR values in **WTP1** and **WTP3**. The median is chosen as the measure of central tendency for the bounded model given its simple closed-form solution. Calculation of mean WTP under this model is more complicated because of the non-linear nature of the error disturbance in the assumed form of WTP (Haab & McConnell, 1997:3,4).

Finally, to examine the effect of potential model mis-specification on the estimates of the welfare measures, the parametric measures of central tendency are compared to those derived from a non-parametric model. Hanemann and Kanninen (1998:402) explain that calculating mean WTP measures using a non-parametric model is useful as this type of model is more robust against possible mis-specification of the response probability distribution and offers the least restricted characterisation of "what the data have to say."[28] This paper calculates mean WTP values using the lower-bound Turnbull method[29] as set out by Haab and McConnell (1996) and Vaughan and Rodriguez (2000). This method uses the proportion of "no" responses at each bid value to provide a discrete stepwise approximation to the WTP cumulative distribution function. As per Table 4, the non-parametric estimates of mean WTP are very similar to those values derived under the linear model for **WTP2** and **WTP4**, suggesting that these models are the least prone to potential mis-specification.

5.3.2 Hypothesis Testing

It can be seen in Section 5.3.1 that calculated WTP values are sensitive to the dependent variable's method of construction. In addition, different functional forms of the WTP models seem to result in different central tendency estimates of household WTP. Mitchell and Carson (1989:194) suggest that where there is no clear statistical or theoretical grounds for preferring one specification over another, reliability requires that the resulting welfare estimates should be similar. This is largely true of the models using the preferred **WTP4** dependent variable, with the exception of the mean calculated under the log (**BIDPR**) model. In order to test the PV hypothesis of Section 4.3, the mean WTP values

[27] Note that the largest sample bid or some other limiting value could have also been used instead of income.
[28] Note that a potential disadvantage of using non-parametric measures is that they only allow limited exploration of the effects of covariates on average WTP measures (Haab & McConnell, 2002:83). This is not a problem in this paper given the simplified models used to estimate WTP values.

derived from the linear **WTP4** models are used. This is because these models display higher goodness-of-fit statistics than the alternative functional forms[30], and appear to have the least mis-specification error upon comparison to the non-parametric estimates. In addition, Crooker and Herriges (2004:468) provide evidence that the simple linear probit specification performs consistently well in estimating mean WTP regardless of the true distribution under consideration.

To test the null hypothesis that the WTP under the two PV scenarios is equal, standard errors for each model are calculated following Cameron and James (1987), and Thaye's t-test is undertaken as detailed in Appendix D. The hypothesis test indicates a calculated t value of 0.1841, which when compared to the critical value of 1.960, suggests that the null hypothesis of an equal mean WTP between the two PVs cannot be rejected. To adjust this hypothesis test to allow for indifferent respondents, this paper follows Haab and McConnell's (2002:136) methodology[31], also outlined in Appendix D. After accounting for those households in the survey that appear indifferent towards the GM labelling program, mean WTP values adjust slightly downwards to £10.95 and £10.65 for the PT and TR models respectively. Using Buckland et al's (1999:113) formula for calculating the standard errors of these new means, the results of the second hypothesis test also indicate that the null hypothesis of equal mean WTP between the two PVs cannot be rejected.

6 – DISCUSSION AND CONCLUSIONS

6.1 Summary of Empirical Findings

The empirical results of this paper are generally consistent with previous biotechnology-related valuation studies, and find that the WTP for GM labelling is affected by price, demographic and attitudinal variables. More specifically, the results indicate a WTP for GM labelling of £10.95 and £10.65 respectively under PT and TR payment mechanisms. These figures, per £100 expenditure on supermarket food, suggest that UK households are WTP an implicit premium of almost 11% for food labelling either through higher food prices or reduced provision of alternative public goods.

Using simple benefits transfer and an estimate of £49,615 million of annual UK supermarket food expenditure (Information Resources, 2004), an approximate yearly benefit of a GM labelling scheme of $5458 million can be calculated. This annual figure easily exceeds an estimated twenty year present value program cost of £1594 million[32] (Jones et al, 2001), suggesting that the implementation of a food labelling program is economically efficient.

---

[29] Other available non-parametric techniques include Kristrom's intermediate-bound measure and Paasche's upper-bound measure. See Vaughan and Rodriguez (2000) for further details.

[30] For example, the X^2 statistics for the linear PT and TR models are 29.2416 and 5.5082 respectively, in comparison to values of 23.3277 and 4.9889 for the bound probit models.

[31] Note that other possible methods for allowing for indifferent respondents include the spike and mixture models. See Buckland et al (1999) for further details and references.

[32] In 2001 prices.

Given that the mean WTP between the two PVs is not statistically significantly different, a food labelling program would be almost equally efficient whether funded out of a new product tax or from existing taxation revenues. The equality between these two WTP measures is an interesting and valuable contribution to the CV literature, because as discussed in the following section, this is not a finding consistent with the previous research in this area.

6.2 PV Discussion and Implications for CV

Indeed, a finding of equal mean WTP between the two PVs is in stark contrast to Bergstrom et al's (2004) findings in their dichotomous-choice survey where the mean WTP under a TR was found to be almost eighteen times higher than that under a standard tax. This is not entirely unexpected, however, as Bergstrom et al (2004:12) explain that their case study results cannot easily be generalized to other populations, regions and public versus private good institutional settings. This is because the relative marginal values of private goods and other public goods are conditional upon existing bundles of these goods, as well as the specific public goods included in the TR payment mechanism. The results of this paper therefore suggest that these factors are likely to be markedly different between the two case studies under consideration.

In the case study investigated here, it would appear that given existing bundles of public and private goods in the UK regions, combined with the preferences of the households living in these areas, respondents are equally willing to trade-off other government-funded goods for the labelling program as they are to trade-off private goods or their own disposable income. That is, CS = CTR, which implies that a researcher estimating WTP for a GM labelling program in these areas could largely rely on the results from a traditional CV study using a PT PV, even if the funding for this program were to actually come from a reallocation of existing taxation funds.

The findings of this paper further suggest that the marginal values of income/private goods, and the marginal value of other public goods are considerably more equal in the UK than they are in the US regions of Maine and Georgia, where the marginal value of income/private goods appears significantly higher. These differences might be due to the US areas having access to a relatively better supply of public goods, or alternatively due to households in the UK regions having relatively higher incomes or larger bundles of private goods. Of course, these differences may not in actual fact be real, as it is likely to be how a household perceives its current bundles of private and public goods that determines how it makes its trade-offs. For example, UK households may not necessarily have relatively larger bundles of private goods than US households, but perhaps may be more content with the bundles they have, thereby displaying relatively lower marginal values for increases in the magnitude of these goods.

It is not clear what impact the choice of policy evaluated might have on the relative values of the alternative welfare measures, given that the two case studies consider different types of programs. For example, it might be that households in the US may be relatively more willing to trade off their own income/private goods if they view a labelling program as more valuable than a ground water quality one. This aspect requires further attention and could be addressed by undertaking studies in different jurisdictions using programs that provide similar benefits to the households in question.

Note that the survey in this paper follows Bergstrom et al's (2004) approach by referring in the TR question to a generic bundle of "other public services," though Bergstrom et al (2004:16) do provide respondents with examples of "roads and bridges, schools, parks, police protection, health care, etc." This may affect the results of our study given that the choice of the bundle of public goods to be traded-off for the labelling program will influence the relative marginal values of public and private goods and therefore the relative magnitudes of CTR and CS.

Bergstrom et al (2004:13) suggest that while using a general bundle of all other public goods may be interpreted somewhat vaguely by respondents, it is effective as it reduces the likelihood of a respondent's valuation being overly influenced or biased by their feelings about one particular public good. However, for more reliable valuation results, despite the possibility of some potential biases, it may be more accurate to use the particular bundle of public goods to be actually traded-off in order to fund the policy in question. That is, if the labelling program in this paper actually were to be funded by existing government revenues, it would be useful for policy evaluation purposes to re-run this survey once the source of these particular funds has been identified.

In conclusion, the TR payment mechanism investigated in this paper presents a number of interesting issues for traditional CV literature. In particular, while our GM case study provides evidence that the welfare measure under a TR is equivalent to that under a more traditional PT PV, this is not supported by Bergstrom et al's (2004) earlier piece of empirical work. Given that results are indeed expected to vary dependent on existing bundles of goods, preferences and the particular public goods included in the TR mechanism, it is of interest to complete further research to determine how these welfare measures are likely to compare over different trade-offs, circumstances and locations.

If findings are not consistent or cannot be a priori predicted, it becomes very important for efficiency reasons that policy evaluation and implementation is carried out using the correct PV context. For example, if a CV evaluation of a particular policy indicates that CS is greater than CTR, with the cost of the policy lying somewhere between the two, implementation of the policy will only be efficient if it is financed out of new taxes, rather than a reduction in existing public goods. Furthermore, even if a policy is efficient when examined under a TR mechanism with a general bundle of public goods, it may not be when evaluated using the specific public good, for example education, which the public funds are actually to be deferred from.

Indeed, if future empirical work continues to confirm that the equivalence of CTR and CS in valuation studies varies by situation, it is essential that future CV studies use the true likely context of payment. Given that a reallocation of taxation funds is likely to be a common means of public policy financing, this suggests that there is likely to be an important role for the tax reallocation payment vehicle in future CV investigations.

APPENDIX A - SURVEY EXCERPT

Product Tax Scenario

New labelling for Non-GM food

Our government is considering a new law that would require all foods are separated into GM and Non-GM food and clearly state their GM content.

For such a programme, the extra cost of separating, transporting, inspecting and labelling Non-GM food is estimated to be **£X for every £100 spent on food by each household.**

The cost of providing people with the chance of being able to distinguish, choose, and afford Non-GM foods would be passed onto the consumer. This would increase the prices of Non-GM foods.

Q1 If you are asked to vote on this program and it would cost your household an extra **£X for every £100 you spend on food** to ensure all foods clearly state their GM content, would you vote yes or no?

☐ Yes, I would be in favour of this programme (Please go to Q2a)

☐ I wouldn't be in favour if the increase is £X but I may support such a programme at some lower amount (Please skip Q2)

☐ No, I'm against this programme (Please go to Q2b)

☐ I don't know (Please skip Q2)

Q2a If Yes, how much food would you consume under the new programme?

☐ As much as usual ☐ 30% less

☐ 5% less ☐ 40% less

☐ 10% less ☐ More than usual

☐ 20% less ☐ Other …………..

Q2b If No, why are you against this programme?

☐ I'm not worried about GM foods ☐ This amount is too high

☐ People who buy Non-GM food should pay the extra price ☐ I need more information

☐ The companies that produce GM food should pay the cost ☐ Other …………………..

☐ The government should pay

Tax Reallocation Scenario

New labelling for Non-GM food

Our government is considering a new law that would require all foods are separated into GM and Non-GM food and clearly state their GM content.

This increases the price of Non-GM foods and would particularly affect *lower income households*.

Some believe that it is the government's obligation to provide all people with the chance of being able to distinguish, choose and afford Non-GM food.

One means of covering this cost would be to use some of the **funds** from the **government purse.** In this case the price of Non-GM foods would **NOT** increase and **NO new taxes** would be introduced. Yet the money used for such inspections could no longer be used to other public services.

Q1 *If you are asked to vote on this programme and it would cost the government* ***£X of your family's taxes for every £100 you spend on food*** *to ensure that all foods clearly state their GM content, would you vote yes or no?*

☐ Yes, I would be in favour of this programme (Please go to Q2a)

☐ I wouldn't be in favour if the increase is £X but I may support such a programme at some lower amount (Please skip Q2)

☐ No, I'm against this programme (Please go to Q2b)

☐ I don't know (Please skip Q2)

Q2a If Yes, how much food would you consume under the new programme?

☐ As much as usual ☐ 30% less

☐ 5% less ☐ 40% less

☐ 10% less ☐ More than usual

☐ 20% less ☐ Other …………..

Q2b If No, why are you against this programme?

☐ I'm not worried about GM foods ☐ This amount is too high

☐ People who buy Non-GM food should pay the extra price ☐ I need more information

☐ The companies that produce GM food should pay the cost ☐ Other …………………..

☐ This money should be spent on other public services

The complete survey, including demographic and attitudinal sections can be found in Kontoleon (2002).

APPENDIX B – SUMMARY STATISTICS AND ESTIMATION RESULTS

TABLE B1: Summary Statistics by Full Sample and PV Treatment

Variable	Full Sample		Standard Tax Treatment		Tax Reallocation Treatment		Thaye's T Statistic
	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	
BIDPR	14.045	11.620	14.614	11.832	13.391	11.359	1.212
GENDER	0.398	0.490	0.417	0.494	0.375	0.485	1.000
AGEMD	52.531	16.551	52.905	16.421	52.100	16.720	0.559
HH	2.632	1.412	2.624	1.392	2.641	1.436	-0.135
CHILD (continuous)	0.580	1.007	0.621	1.066	0.532	0.935	1.015
EDU (continuous)	3.281	1.305	3.263	1.331	3.302	1.276	-0.346
INCOME	2906.868	2950.472	3065.954	3123.697	2724.046	2732.754	1.335
F1	-0.004	0.978	0.009	0.977	0.002	0.981	0.0958
F2	0.006	0.912	0.011	0.922	-0.001	0.901	0.166
F3	0.012	0.869	0.023	0.865	-0.001	0.873	0.327
F4	-0.004	0.841	0.000	0.803	-0.008	0.883	0.118
F5	-0.002	0.820	0.001	0.844	-0.005	0.793	0.084

Under a critical T-value of 1.96, a null hypothesis of equal means across the treatments is not rejected.

TABLE B2: Comparison of Mean Values by Group

Variable	Households included in WTP4	Households excluded from WTP4 (for protest / indifference)	Households excluded from original sample
GENDER	0.382	0.438	0.444
AGEMD	52.406	52.843	65.175
HH	2.629	2.641	2.068
CHILD (continuous)	0.542	0.673	0.327
EDU (continuous)	3.345	3.229	3.100
INCOME	2994.834	2668.389	1978.667
F1	0.059	-0.163	0.010
F2	0.042	-0.083	-0.129
F3	0.062	-0.111	-0.268
F4	0.028	-0.082	0.090
F5	-0.076	0.182	0.039

TABLE B3: Initial Estimation Results using WTP4

VARIABLE	ST Marginal Effect	P-value	TR Marginal Effect	P-value
Constant	-0.2110	0.2983	0.0005	0.9981
BIDPR	-0.0195	0.0000**	-0.0094	0.0082**
GENDER	-0.0132	0.8405	0.0650	0.4486
AGE	0.0074	0.0125**	0.0006	0.8389
EDU	-0.0475	0.5482	-0.0941	0.2703
INCOME	0.0000	0.4032	0.0000	0.5487
CHILD	-0.0112	0.9197	-0.0537	0.6506
HHOLD	-0.0469	0.6608	0.1870	0.1125
F1	0.0813	0.0441**	0.1690	0.0001**
F2	0.1122	0.0138**	0.0656	0.1515
F3	0.0281	0.5484	0.0940	0.0531*
F4	-0.0925	0.0664*	0.0624	0.2240
F5	-0.0887	0.0855*	0.0201	0.7076
Pseudo R^2 Test sig ~X^2; Prob > X^2 Prediction Success	19.69% 52.2900; 0.0000 72.91%		12.76% 31.2961; 0.0018 70.06%	

** Significant at the 5% level; * significant at the 10% level

APPENDIX C – CLASSIFICATION TABLES AND LIMDEP COMMAND FILE

Table C1: Classification Table: PT Model

Actual	Predicted		TOTAL
	0	1	
0	81	26	107
1	28	68	96
TOTAL	109	94	203

PR (Actual = 0| Predicted = 0) = 74.31%
PR (Actual = 1| Predicted = 1) = 72.34%
PR (Predicted = 0| Actual = 0) = 75.70%
PR (Predicted = 1| Actual = 1) = 70.83%
Correctly classified = **73.40%**

Table C2: Classification Table: TR Model

Actual	Predicted		TOTAL
	0	1	
0	61	29	90
1	31	56	87
TOTAL	92	85	177

PR (Actual = 0| Predicted = 0) = 66.30%
PR (Actual = 1| Predicted = 1) = 65.88%
PR (Predicted = 0| Actual = 0) = 67.78%
PR (Predicted = 1| Actual = 1) = 64.36%
Correctly classified = **66.10%**

LIMDEP Command File Excerpt

```
/* Create New Variables */
CREATE; LINCOME=LOG(INCOME); AGESQ=AGE*AGE; LBIDPR=LOG(BIDPR); BOUND=
LOG((INCOME-BIDPR)/BIDPR) $
/* Summary Statistics */
SAMPLE; ALL $
DSTAT; RHS=BIDPR,GENDER,AGE,HH,CHILD,EDUDB,INCOME,F1,F2,F3,F4,F5 $
SAMPLE; 1-285 $
DSTAT; RHS=BIDPR,GENDER,AGE,HH,CHILD,EDUDB,INCOME,F1,F2,F3,F4,F5 $
SAMPLE; 286-533 $
DSTAT; RHS=BIDPR,GENDER,AGE,HH,CHILD,EDUDB,INCOME,F1,F2,F3,F4,F5 $
/* Final Estimation with WTP4 */
SAMPLE; 1-285 $
PROBIT; LHS=WTP4; RHS=ONE,BIDPR,GENDER,AGE,AGESQ,EDU,LINCOME,CHILD,
HHOLD,F1,F2,F3,F4,F5  ; MARGINAL EFFECTS $
SAMPLE; 286-533 $
PROBIT; LHS=WTP4; RHS=ONE,BIDPR,GENDER,AGE,AGESQ,EDU,LINCOME,CHILD,
HHOLD,F1,F2,F3,F4,F5  ; MARGINAL EFFECTS $
/* Initial Step of WTP Calculations with WTP4 */
SAMPLE; 1-285 $ PROBIT; LHS=WTP4; RHS=ONE,BIDPR $
SAMPLE; 286-533 $ PROBIT; LHS=WTP4; RHS=ONE,BIDPR $
SAMPLE; 1-285 $ PROBIT; LHS=WTP4; RHS=ONE,LBIDPR $
SAMPLE; 286-533 $ PROBIT ; LHS=WTP4; RHS=ONE,LBIDPR $
SAMPLE; 1-285 $ PROBIT; LHS = WTP4 ; RHS = ONE, BOUND $
SAMPLE; 286-533 $ PROBIT; LHS = WTP4 ; RHS = ONE, BOUND $
```

APPENDIX D – HYPOTHESIS TESTING

PV Hypothesis Test 1

H_0: $WTP_{PT} = WTP_{TR}$; H_1: $WTP_{PT} \neq WTP_{TR}$

$\mu_{PT} = 11.98$; $S_{PT} = 2.14$; $N_{PT} = 203$

$\mu_{TR} = 11.91$; $S_{TR} = 4.91$; $N_{TR} = 177$

$$T = \frac{(\mu_{PT} - \mu_{TR})}{S_P \sqrt{1/N_{PT} + 1/N_{TR}}}$$

T has a student's t distribution with $N_{PT} + N_{TR} - 2$ (378) degrees of freedom, and a critical value of approximately 1.96 under a 5% level of significance two-tailed t test.

$$S_P^2 = \frac{[(N_{PT}-1)S_{PT}^2 + (N_{TR}-1)S_{TR}^2]}{(N_{PT} + N_{TR} - 2)}$$

$$S_P^2 = \frac{[(202)(2.14)^2 + (176)(4.91)^2]}{(203 + 177 - 2)} = 13.6722$$

$$S_P = 3.6976$$

$$T = \frac{(11.98 - 11.91)}{(3.6976)(\sqrt{1/203 + 1/177})} = 0.1841$$

As T does not exceed the critical value, the null hypothesis that the mean WTP under each PV is equal cannot be rejected.

PV Hypothesis Test 2 – Allowing for Indifference

H_0: $WTP_{PTIND} = WTP_{TRIND}$; H_1: $WTP_{PTIND} \neq WTP_{TRIND}$

$\mu_{PT} = 11.98$; $S_{PT} = 2.14$; $V_{PT} = 4.5796$; $N_{PT} = 203$; $N_{PTIND} = 19$; $N_{TOTPT} = 222$

$\mu_{TR} = 11.91$; $S_{TR} = 4.91$; $V_{TR} = 24.1081$; $N_{TR} = 177$; $N_{PTIND} = 21$; $N_{TOTTR} = 198$

$$\pi_{PT} = N_{PTIND} / N_{TOTPT} = 0.0856$$

$$\pi_{TR} = N_{TRIND} / N_{TOTTR} = 0.1061$$

$$WTP_{PTIND} = WTP_{PT}(1 - \pi_{PT}) = 10.9545$$

$$WTP_{TRIND} = WTP_{TR}(1 - \pi_{TR}) = 10.6463$$

$$Var(\pi_{PT}) = \frac{\pi_{PT}(1 - \pi_{PT})}{N_{TOTPT}} = 0.00035$$

$$Var(\pi_{TR}) = \frac{\pi_{TR}(1 - \pi_{TR})}{N_{TOTTR}} = 0.00048$$

$$Var(WTP_{PTIND}) = (1 - \pi_{PT})^2 \operatorname{var}(WTP_{PT}) + WTP_{PT}{}^2 \operatorname{var}(\pi_{PT}) + \operatorname{var}(WTP_{PT}) \operatorname{var}(\pi_{PT}) = 3.8809$$

$$S_{PTIND} = \sqrt{Var(WTP_{PTIND})} = 1.97$$

$$Var(WTP_{TRIND}) = (1 - \pi_{TR})^2 \operatorname{var}(WTP_{TR}) + WTP_{TR}{}^2 \operatorname{var}(\pi_{TR}) + \operatorname{var}(WTP_{TR}) \operatorname{var}(\pi_{TR}) = 23.9164$$

$$S_{TRIND} = \sqrt{Var(WTP_{TRIND})} = 4.8904$$

$$S_P^2 = \frac{[(221)(1.97)^2 + (197)(4.8904)^2]}{(222 + 198 - 2)} = 13.3233$$

$$S_P = 3.6501$$

$$T = \frac{(10.9545 - 10.6463)}{(3.6501)(\sqrt{1/222 + 1/198})} = 0.8638$$

As T in this second hypothesis test does not exceed the critical value, the null hypothesis that the mean WTP under each PV is equal again cannot be rejected.

REFERENCE LIST

Aldrich, J.H. & Nelson, F.D. 1984. *Linear Probability, Logit and Probit Models*. Newbury Park: Sage Publications.

Arrow, K., Solow, R., Portney, P.P., Leamer, E.E., Radner, R. & Schuman, H. 1993. 'Report on the National Oceanic and Atmospheric Administration Panel on Contingent Valuation.' *Federal Register*, 58 (10): 4602-4614.

Bateman, I.J., Langford, I.H., Turner, R.K., Willis, K.G. & Garrod, G.D. 1995. 'Elicitation and truncation effects in contingent valuation studies.' *Ecological Economics*, 12: 161-179.

Bateman, I. (ed), Carson, R., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce, D., Sugden, R., Sawanson, J. 2002. *Economic Valuation with Stated Preference Tech*niques (In Association With the Dtlr and Defra). Cheltenham: Edward Elgar Publishing.
(NB: Draft Page numbers used)

Bergstrom, J.C., Boyle, K.J. & Yabe, M. 2004. 'Trading Taxes vs, Paying Taxes to Value and Finance Public Environmental Goods.' Forthcoming in *Environmental and Resource Economics.*

Boccaletti, S. & Moro, D. 2000. 'Consumer Willingness-To-Pay for GM Food Products in Italy.' *AgBioForum*, 3(4): 259-267.

Bredahl, L. 2000. 'Determinants of Consumer Attitudes and Purchase Intentions with Regard to Genetically Modified Foods – Results of a Cross-National Survey.' The Aarhus School of Business, Working Paper No. 69.

Buckland, S.T., Macmillan, D.C., Duff, E.I. & Hanley, N. 1999. 'Estimating Mean Willingness to Pay from Dichotomous Choice Contingent Valuation Studies.' *The Statistician,* 48(1): 109-124.

Burton, M. & Pearse, D. 2002. 'Consumer Attitudes Towards Genetic Modification, Functional Foods, and Microorganisms: A Choice Modeling Experiment for Beer.' *AgBioForum*, 5(2): 51-58.

Cabinet Office Strategy Unit. July 2003. 'Field Work: weighing up the costs and benefits of GM crops.' http://www.number- 10.gov.uk/su/gm/downloads/gm_crop_report.pdf

Cameron, T. & James, M.D. 1987. 'Efficient Estimation Methods for "Close-Ended" Contingent Valuation Surveys.' *The Review of Economics and Statistics*. 69: 269-276.

Carson, R.T., Hanemann, W.M., Kopp, R.J., Krosncik, J.A., Mitchell, R.C., Presser, S, Ruud, P.A., Smith, K.V., Conaway, M. & Martin, K. 1998. 'Referendum design and contingent valuation: The NOAA panel's No-Vote recommendation.' *The Review of Economics and Statistics*, 80: 335-358.

Carson, R.T. 1997. 'Contingent Valuation: Theoretical Advances and Empirical Tests Since the NOAA Panel.' *American Journal of Agricultural Economics*, 79:1501-1507.

Carter, C.A. & Gruère, G.P. 2003. 'Mandatory Labelling of Genetically Modified Foods: Does it Really Provide Consumer Choice?' Working Paper, Department of Agricultural and Resource Economics, University of California, Davis.

Champ, P.A., Flores, N.E., Brown, T.C. & Chivers, J. 2002. 'Contingent Valuation and Incentives.' *Land Economics*, 78(4): 591-604.

Chern, W.S. & Kaneko, N. 2003. 'Contingent Valuation of the Willingness to Pay for Selected Genetically Modified Foods: A Multi-Country Analysis.' 7th International ICABR Conference, Ravello, Italy, June 29 - July 3.

Chern, W.S., Rickertsen, K., Tsuboi, N & Fu, T. 2002. 'Consumer Acceptance and Willingness to Pay for Genetically Modified Vegetable Oil and Salmon: A Multiple-Country Assessment.' *AgBioForum*, 5(3): 105-112.

Crooker, J.R. & Herriges, J.A. 2004. 'Parametric and Semi-Nonparametric Estimation of Willingness-to-Pay in the Dichotomous Choice Contingent Valuation Framework.' *Environmental and Resource Economics*, 27: 451-480.

Cummings, R.G, Brookshire, D. & Schulz, W.D. (ed.s).1986. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Methodology*. Savage: Rowman & Littlefield Publishers.

Diamond, P.A. & Hausman, J.A. 1994. 'Contingent Valuation: Is Some Number Better than No Number?' *Journal of Economic Perspectives*, 8(4): 45-64.

European Commission. 2000. 'Economic Impacts of Genetically Modified Crops on the Agri-Food Sector: A First Review.' Brussels: Directorate-General for Agriculture.

Fomby, T.B., Hill, R. & Stanley, R. 1988. *Advanced Econometric Methods*. New York: Springer-Verlag.

Food Standards Agency. 2003. 'Date set for new GM food regulation.' Thursday, 06 November. http://www.foodstandards.gov.uk/news/newsarchive/new_reg

Gamble, J., Muggleston, S., Hedderley, D., Parminter, T. & Richardson-Harman, N. 2002. *Genetic Engineering: The Public's Point of View*. Auckland: HortResearch.

Garrod, G. & Willis, K.G. 1999. *Economic Valuation of the Environment: Methods and Case Studies*. Cheltenham: Edward Elgar Publishing.

Gaskell, G. & Allum, N. 2003. 'Europeans and Biotechnology in 0220: Eurobarometer 58.0.' DG Research. European Commission.

Green, C.H. & Tunstall, S.M. 1999. 'Contingent Valuation: a Psychological Perspective' In Bateman, I. & Willis, K. (eds). *Contingent Valuation: European, North American and World Experience*, Oxford: Oxford University Press.

Greene, W. 1995. *LIMDEP Version 7.0 Reference Guide*. New York: Econometric Software Inc.

Greene, W.H. 1993. *Econometric Analysis* (2nd ed). Englewood Cliffs: Prentice Hall.

Greenley, D.A., Walsh, R.G. & Young, R.A. 1981. 'Option Value: empirical evidence from a case study of recreation and water quality.' *Quarterly Journal of Economics*, 96 (4): 657-672.

Grimsrud, K.M., McCluskey, J.J., Loureiro, M.L. & Wahl, T.I. 2003. 'Consumer Attitudes toward Genetically Modified food in Norway.' IMPACT Center Working Paper.

Groothuis, P.A. & Whitehead, J.C. 2002. 'Does don't know mean no? Analysis of 'don't know' responses in dichotomous choice contingent valuation questions.' *Applied Economics*, 34: 1935-1940.

Gujarati, D.N. 1995. *Basic Econometrics* (3rd ed). New York: McGraw Hill.

Haab, T.C.Y. & McConnell, K.E. 2002. *Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation* (New Horizons in Environmental Economics). Cheltenham: Edward Elgar Publishing.

Haab, T.C.Y. & McConnell, K.E. 1997. 'A Simple Method for Bounding Willingness to Pay Using a Probit or Logit Model.' Working Paper 9713, Department of Economics, East Carolina University.

Haab, T.C.Y. & McConnell, K.E. 1996. 'Referendum Models and Negative Willingness to Pay: Alternative Solutions.' Working Paper 9610, Department of Economics, East Carolina University.

Hanemann, W.M. 1995. 'CV and Economics.' In Willis, K.G. & Corkindale, J.T. (ed.s) *Environmental Valuation: New Perspectives*. Wallingford: CAB International.

Hanemann, W.M. & Kanninen, B. 1998. 'The Statistical Analysis of Discrete-Response CV Data.' Department of Agricultural and Resource Economics and Policy Division of Agricultural and Natural Resources, University of California at Berkeley, Working Paper No. 798.

Hanley, N., Shogren, J.F. & White, B. 1997. *Environmental economics in theory and practice.* Basingstoke: Macmillan.

Hoevenagel, R.1990.'The validity of the Contingent Valuation Method: Some aspects on the bias of three Dutch studies.' Paper presented at the E.A.E.E. Congress in Venezia, Italy, April 17-20.

Huffman. W.E., Shogren, J.F., Rousu, M. & Tegene, A. 2001.'The Value to Consumers of GM Food Labels in a Market with Diverse Information: Evidence from Experimental Auctions.' Iowa State University, Department of Economics. Staff Paper #345.

Information Resources, 2004. "Supermarket Review, 26th April." www.infores.com/public/uk/prodserv/uk_prd_superreview.htm

Jakobsson, K.K. & Dragun, A.K. 1996. *Contingent Valuation and Endangered Species: Methodological Issues and Applications*. Cheltenham: Edward Elgar Publishing.

Johansson, M.V. 1999. 'Valuing a Peripheral Environmental Amenity. The Swede's Willingness to Pay for the Survival of the African Elephant.' Working Paper, Department of Economics, Umea University.

Johnston, R.J., Swallow, S.K., & Weaver, T.F.1999. 'Estimating Willingness to Pay and Resource Trade-offs with Different Payment Mechanisms: An Evaluation of a Funding Guarantee for Watershed Management.' *Journal of Environmental Economics and Management*, 38: 97-120.

Jones, S., Bramley-Harker, E., Aslam, S., Young, J., Spackman, M. & Dodgson, J. 2001. 'Economic Appraisal of Options for the Extension of Legislation of GM Labelling.' A Final Report for the Food Standards Agency. London: National Economic Research Associates.

Kaneko, N. & Chern, W.S. 2003. 'Consumer Acceptance of Genetically Modified Foods: A Telephone Survey.' *Consumer Interests Annual*, 49: 1-13.

Kato, T & Hidano, N. 2002. 'An Empirical Comparison between Tax Payment and Donation in a Contingent Valuation Survey: Value of Preserving the Satsunai River, Japan.' Discussion Paper 02-04, Department of Social Engineering, Tokyo Institute of Technology.

Kaye-Blake, W., Bicknell, K. & Lamb, C. 2002. 'Willingness to Pay for Genetically Modified Food Labelling in New Zealand.' 6th International ICABR Conference

Ravello, Italy, July 11-14.

Kontoleon, A. 2002. 'Essays on non-market valuation of environmental resources: policy and technical explorations.' PhD Thesis, Department of Economics, University College London.

Kristrom, B. 1990. 'A non-parametric approach to the estimation of welfare measures in discrete response valuation studies.' *Land Economics*, 66: 135-139.

Larson, H.J. 1982. *Introduction to Probability Theory and Statistical Inference* (3rd ed.). New York: Wiley.

Lusk, J., Daniel, M.S., Mark, D. & Lusk, C. 2001. 'Alternative Calibration and Auction Institutions for Predicting Consumer Willingness to Pay of Nongenetically Modified Corn Chips.' *Journal of Agricultural and Resource Economics,* 26:40-57.

Maddala, G.S. 1983. *Limited-dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.

Mendenhall, C.A. & Evenson, R.E. 2002. 'Estimates of Willingness to Pay a Premium for Non-GM Foods: a Survey.' In Santaniello, V., Evenson, R.E. & Zilberman, D. 2002. *Market Development for Genetically Modified Foods*. CAB International.

McCluskey, J.J., Ouchi, H., Grimsrud, K.M. & Wahl, T.I.. 2003. 'Consumer Response to Genetically Modified Food Products in Japan.' *Agricultural and Resource Economics Review* 32(2*).*

Mitchell, R.C. and R.T. Carson.1989. *Using Surveys to Value Public Goods: the Contingent Valuation Method*. Washington: john Hopkins University Press for Resources for the Future.

Moon, W. & Balasubramanian, S.K. 2001. 'Public Perceptions and Willingness-To-Pay a Premium for Non-GM Foods in the US and UK.' *AgBioForum*, 4(3,4): 221-231.

Morrisson, M.D., Blamey, R.K. & Bennett, J.W. 2000. 'Minimising Payment Vehicle Bias in Contingent Valuation Studies.' *Environmental and Resource Economics*, 16: 407-422.

Noussair, C., Robin, S. & Ruffieux, B. 2002. 'Consumer behavior with regard to genetically modified organisms in the food supply.' http://www.emory.edu/COLLEGE/ECON/faculty/noussair/.

Nunes, P.A.L.D. 2002. 'Using factor analysis to identify consumer preferences for the protection of a natural area in Portugal.' *European Journal of Operational Research*, 140: 499-516.

Pearce, D. 2003. 'Estimating the benefits of Sewer Flooding Control.' Report for the National Audit Office. http://www.nao.org.uk/publications/nao_reports/03-04/ofwat_sewerage_study.pdf

Phillips, Peter W.B. and Heather McNeill. 2000. "A Survey of National Labelling Policies for GM Foods." *AgBioForum*, 3(4): 219-224.

Rowe, R.D., d'Arge, R. & Brookshire, D. 1980. 'An Experiment on the Economic Value of Visibility.' *Journal of Environmental Economics and Management*, 7: 1-19.

Russell, C.S. 2001. *Applying Economics to the Environment*. New York: Oxford University Press.

Sheridan, B. 2001. *EU biotechnology law and practice: regulating genetically modified and novel food products*. Bembridge, Isle of Wight: Palladian Law Publishing.

Teisl, M.F., Roe, B., Vayda, M. & Ross, N. 2003. 'Willingness to Pay for Genetically Modified Foods with Bundled Health and Environmental Attributes.' 7th International ICABR Conference, Ravello, Italy, June 29 - July 3.

Vaughan, W.J., Russell, C.S., Rodriguez, D.J. & Darling, A.H. 1999. 'Willingness to Pay: Referendum Contingent Valuation and Uncertain Project Benefits.' Washington, D.C: Inter-American Development Bank, Sustainable Development Department Technical Papers Series, 130.

Vaughan, W.J. & Rodriguez, D.J. 2000. 'A Note on Obtaining Welfare Bounds in Referendum Contingent Valuation Studies.' Washington, D.C: Inter-American Development Bank, Sustainable Development Department, Environment Division.

Wang, H. 1997. 'Treatment of "Don't-Know" Responses in Contingent Valuation Surveys: A Random Valuation Model.' *Journal of Environmental Economics and Management*, 32: 219-232.