

The benefits of farm animal welfare legislation: the case of the EU broiler directive and truthful reporting

Article

Accepted Version

Bennett, R. ORCID: <https://orcid.org/0000-0003-3226-8370>, Balcombe, K., Jones, P. ORCID: <https://orcid.org/0000-0003-3464-5424> and Butterworth, A. (2019) The benefits of farm animal welfare legislation: the case of the EU broiler directive and truthful reporting. *Journal of Agricultural Economics*, 70 (1). pp. 135-152. ISSN 0021-857X doi: <https://doi.org/10.1111/1477-9552.12278> Available at <https://centaur.reading.ac.uk/75872/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1111/1477-9552.12278>

Publisher: Wiley-Blackwell

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

The benefits of farm animal welfare legislation: the case of the EU broiler directive and truthful reporting

Richard Bennett, Kelvin Balcombe, Philip Jones and Andrew Butterworth¹

[Original submitted October 2016, Revisions received February 2017 to January 2018, Accepted February 2018]

ABSTRACT

The EU broiler Directive came into force in the UK in June 2010 with the aim of setting new minimum standards, monitoring broiler welfare and addressing any welfare problems. A survey questionnaire was used to elicit information from a stratified sample of citizens in England and Wales regarding their willingness to pay for the provisions of the Directive, as an estimate of the consumer surplus associated with the legislation. We also explore the usefulness of Prelec's (2004) Bayesian Truth Serum (BTS) in promoting respondents' truthful reporting. A median willingness to pay of £21.5 per household per year (corrected for sample bias and possible 'yea saying') was estimated from 665 responses. This provides an estimated benefit of the legislation to citizens of over £503 million/yr, equivalent to 5.3% of current consumer expenditure on chicken. This compares to an estimated £22 million per annum cost of producers' compliance and government enforcement associated with the legislation. No statistically significant differences in responses between respondents that did and did not have a BTS incentive to answer questions truthfully were found, which might reflect apparently truthful answers in this case, an insufficiently strong financial incentive or a weakened effect due to an element of disbelief in the BTS amongst the sample. The analysis suggests that the benefits of the broiler Directive to citizens greatly outweigh the additional costs to producers, making a case for the legislation to be retained.

¹ Richard Bennett (contact author: E-mail r.m.bennett@reading.ac.uk) Kelvin Balcombe and Philip Jones are all in the School of Agriculture, Policy and Development, University of Reading, UK. Andrew Butterworth is in the Department of Clinical Veterinary Science, University of Bristol, UK. The authors would like to thank the Department for Environment, Food and Rural Affairs for funding the study to evaluate the effectiveness of implementation of the EU Broiler Directive in England and Wales, of which the research reported here formed a part. We would also like to thank the producers and citizens who participated in the study. In addition, our thanks are due to two anonymous reviewers and the Editor for comments on earlier drafts of this paper. Any errors or omissions are the responsibility of the authors.

- 25 **KEYWORDS:** Animal welfare; EU broiler legislation; willingness to pay; truthful reporting;
- 26 Bayesian Truth Serum
- 27 **JEL Classifications:** *C11, C18, D12, D46*

28 **The benefits of farm animal welfare legislation: the case of the EU broiler** 29 **directive and truthful reporting**

30 **1. Introduction**

31 Legislation has been a major policy instrument for the protection of animal welfare in Europe with
32 many European Union (EU) regulations and directives implemented in the last 20 years (Bennett
33 and Appleby, 2011; European Commission, 2016). In 2007 new EU rules were agreed (Council
34 Directive 2007/43/EC) for protecting the welfare of broiler chickens (European Commission,
35 2007). The broiler Directive came into force in the United Kingdom (UK) on 30 June 2010. When
36 such legislation is introduced there is an expectation that the benefits of that legislation will be
37 greater than the costs. Expected costs include costs to government, for administration, monitoring
38 and enforcement, as well as compliance costs for producers and broader industry and national costs
39 associated with impacts on competitiveness, international trade and economic output. Expected
40 benefits of such legislation are generally non-monetary, but include consumer perceptions of
41 enhanced chicken meat quality and enhancement of consumer and citizens levels of satisfaction
42 from knowing that the welfare of broiler chickens is better protected in the food production process
43 (see Mayfield *et al.*, 2007 in relation to consumer attitudes to animal welfare in Europe). Indeed,
44 a UK government *ex ante* Impact Assessment (IA) published in 2010 (Defra, 2010) concluded that
45 while the broiler Directive would lead to some costs to producers, this would be more than off-set
46 by non-market benefits to consumers, in the form of higher broiler welfare, for which many would
47 be willing to pay. We test the hypothesis that citizens have an additional willingness to pay (WTP)
48 for the broiler Directive beyond any price increase as a result of increased costs of production due
49 to the legislation and assess the scale of such benefits by means of an *ex post* WTP survey of
50 citizens. Since the perceived costs of improving animal welfare can deter legislation it is important
51 to estimate the benefits of legislation to better inform policy. In addition, as a methodological
52 contribution, we explore the usefulness of Prelec's (2004) Bayesian Truth Serum (BTS) which is
53 designed to reduce respondents' misreporting (i.e. help ensure 'truthful' reporting) in surveys.

54
55 Theoretically, the WTP we estimate using the CV method is a Hicksian consumer surplus measure
56 (see Mitchell and Carson, 1989 p25) which can be interpreted as a respondent's maximum WTP
57 for the broiler legislation and its provisions (considered, at least in part, as a public good, because
58 some of the benefits accruing from the legislation are non-excludable and non-divisible/non-rival).

59 Our WTP estimate measures both use and non-use value and, as noted by Zhao and Kling (2004),
60 is defensible theoretically as a welfare measure that can be directly applied to cost-benefit analysis.

61
62 Section 2 covers the background of WTP in the area of animal health and welfare and associated
63 legislation along with a description of the broiler Directive and its implementation in the United
64 Kingdom. Section 3 outlines the WTP survey whilst Section 4 covers the analytical methods.
65 Section 5 presents the results and Section 6 discusses them. Section 7 offers some conclusions.

66

67 **2. Background**

68

69 **WTP applied to animal health and welfare**

70 Although widely used in other areas, stated preference valuation methods have perhaps been less
71 widely applied to animal welfare. Bennett *et al.* (2011) provide a review and critique of valuation
72 studies applied to animal welfare. Lagerkvist and Hess (2011) identify 24 stated preference
73 studies of people's WTP in relation to animal welfare, yielding 106 WTP estimates, in their
74 meta-analysis. Approximately half of these are contingent valuation studies and half use a
75 choice experiment approach (with one other using an experimental auction method).

76 Additionally, Clark *et al.* (2016) observed just 17 WTP studies out of some 80 studies included
77 in their systematic review of public attitudes, perceptions and behaviours towards animal welfare
78 concerns arising from livestock production diseases.

79

80 A common feature of these studies is that they elicit WTP from citizens regarding specific
81 changes in husbandry or other practices to improve welfare, although Bennett *et al.* (2012)
82 present a method for the economic valuation of animal welfare benefits more generally using a
83 single welfare score. WTP values can be used as estimates of the likely magnitude of the benefits
84 that citizens obtain from each of these animal husbandry or other measures. It is clear from a
85 number of the studies reviewed that consumers value animal welfare not only for ethical reasons
86 but also in some cases because they believe that products from animals with higher welfare are
87 of higher quality in terms of taste, nutrition and safety and better for the environment (despite
88 there being little empirical evidence to support the validity of these beliefs).

89

90 **The broiler Directive and its implementation in the UK**

91 While one of the purposes of the broiler Directive is to promote more universal achievement of
92 minimum animal welfare standards across EU members states (MS), a number of
93 implementation options are permitted, including variation in the maximum stocking density
94 limits that can be applied in each MS. In the UK producers have been permitted to choose from
95 among two stocking density maxima:

96 i) stocking up to 33kg live weight per m² is permitted if specific standards are met for drinkers,
97 feeding, litter, ventilation and heating, noise, light, inspections, cleaning, record keeping, training
98 and surgical interventions.

99 ii) stocking beyond 33kg per m² up to 38kg/m² is permitted if an additional set of standards are
100 met (the Annex II requirements). These include notification and documentation requirements,
101 plus further controls on environmental parameters in broiler housing.

102

103 In addition, the Directive requires the collection of data from farms on cumulative daily
104 mortality (CDM) and data from slaughterhouses for eight post-mortem measures of body
105 condition (collectively known as the ‘trigger conditions’) to help identify poor welfare on
106 farms. These data, which relate to each batch of birds per farm sent to slaughter, are used
107 by the Food Standards Agency (FSA) and the Animal and Plant Health Agency (APHA) to
108 identify farms which may require problem notification and/or on-site inspection. The post-
109 mortem body condition measures for which data are collected are: 1) ascites/oedema, 2)
110 cellulitis and dermatitis, 3) dead on arrival, 4) emaciation, 5) joint lesions/arthritis, 6)
111 septicaemia/respiratory problems, 7) total carcass rejections and 8) foot pad dermatitis
112 score. Evaluation of these data involves two processes:

113 Process 1: An alert (to APHA) is triggered if the incidence of any of the individual post-
114 mortem conditions is exceptionally high in any batch (defined as greater than six standard
115 deviations above the mean);

116 Process 2: An alert is triggered if the CDM is unusually high (defined as greater than three
117 standard deviations above the mean) and, additionally, the level of three or more of the
118 post-mortem conditions is high (defined as above the mean).

119

120 When trigger thresholds are breached, the keeper of the animals and APHA are alerted by

121 means of a ‘trigger report’. APHA and FSA have inspection regimes and data handling
122 systems to communicate information relating to poor welfare between the slaughterhouse
123 and the producer. Investigative action will be taken by APHA Veterinary Officers, and this
124 may include requesting a written action plan to remedy the problem from producers and/or
125 a visit to the production site. APHA may, in addition, carry out a number of random
126 welfare inspections.

127
128 Thus, under the broiler Directive, new broiler production standards are set and enforced
129 through a system which continuously monitors key welfare indicators, linked to an
130 intervention process to deal with problems should they occur. In 2013/14 the percentage of
131 batches sent for slaughter in GB which exceeded Trigger 1 levels varied from less than
132 0.2% for joint and arthritis problems to 1.8% for foot pad dermatitis (Food Standards
133 Agency, personal communication).

134

135 **3. The Survey**

136 **Questionnaire design and survey administration**

137 A questionnaire was designed which contained an introduction explaining the nature of the
138 survey and its purpose, followed by questions to respondents with regards to their:

- 139 1. Personal characteristics (sample stratification variables);
- 140 2. Current consumption of chicken meat;
- 141 3. Attitudes towards farmed-animal welfare;
- 142 4. WTP for farmed-animal welfare improvements in general;
- 143 5. Attitudes towards the broiler Directive;
- 144 6. WTP for the Directive and debriefing questions to explain respondents’ WTP responses;
- 145 7. Socio-demographic characteristics (non-stratification variables).

146

147 A specimen copy of the questionnaire can be found in the online Appendix, available at the
148 publisher’s website. Two versions of the questionnaire were designed in relation to those who did,
149 and did not, consume chicken. Chicken consumers were asked how much extra they would be
150 willing to pay for the Directive in the form of a premium on the price of the conventionally reared
151 chicken products that they purchased, while non-chicken eaters were asked their WTP as an

152 additional sum on their income tax. The survey was carried out by means of a web-based
153 questionnaire hosted by the commercial research support company Qualtrics
154 (<http://www.qualtrics.com>). The questionnaire was sent to a Qualtrics-secured panel of the general
155 public in England and Wales. A stratification procedure was applied during recruitment of survey
156 participants to ensure that the sample was broadly representative of populations in England and
157 Wales. This required set proportions across geographic area, age and income, based on population
158 distributions from data from the Office of National Statistics (ONS).² Respondents were screened
159 to ensure that they were the primary, or regular, purchaser of food for their household.

160
161 Respondent WTP was elicited using a contingent valuation (CV) payment card with a discrete
162 dichotomous choice format with multiple increasing values akin to a ‘bidding game’ (see, for
163 example, Heinzen and Bridges, 2008, who compare four different CV elicitation methods). The
164 payment card method is regarded as efficient, robust and reliable (see, for example, Bateman, *et*
165 *al*, 2002; Pearce and Ozdemiroglu, 2002), and was considered the most appropriate WTP
166 elicitation method for our purposes. We sought people’s holistic WTP for a single specified
167 legislative intervention rather than valuation of different attributes of the legislation where a choice
168 experiment approach would have been more appropriate. Our hypothesis was that people would
169 have a willingness to pay in addition to current market prices of chicken meat to retain the
170 legislation up to the point where their consumer surplus is zero. After reminding respondents that
171 they have a limited household budget and that additional money spent on supporting the chicken
172 legislation may mean that they have less money to spend on other things, respondents were
173 presented with a range of seven bids, of ascending value, from ‘5 pence per week extra’ to ‘more
174 than £4 per week extra’. These bids were expressed as an additional amount that respondents would
175 pay per week, either in the form of an increase in the price of chicken meat, or taxation for those
176 relative few who did not consume chicken³. The range of amounts chosen was based on findings

² Despite this, some over-recruitment was found in certain socio-demographic groups requiring adjustment described later in the Results section. The questionnaire was pre-tested by the research team on a small number of consumers, and then formally piloted by Qualtrics using a panel of 50 consumers. The survey was officially launched on 30 September 2014 and was closed on 8 October 2014, by which time 665 usable responses had been received. In order to eliminate the possibility of partially completed questionnaires, respondents were required to complete all questions before they could file their return.

³ A zero amount was not presented to respondents at this point because (i) respondents could answer ‘No’ to all the bids presented to them including the lowest 5 pence per week bid and (ii) a previous question had already

177 of pre-testing of the questionnaire and was later confirmed to be appropriate in the pilot.
178 Respondents were asked to tick ‘Yes’ if they would be willing to pay each amount or ‘No’ if they
179 would not. Respondents were asked to state their WTP to support the Directive and its associated
180 provisions, taken as a whole. Respondents were also asked to indicate how they thought other
181 people would respond to the WTP questions.

182

183 **Misreporting and the Bayesian Truth Serum (BTS)**

184 Prelec (2004) proposed the BTS as a general mechanism to encourage ‘truth telling’ in a variety
185 of contexts, from answering simple ‘yes or no’ questions to more complex WTP studies. There
186 has been growing recognition of the BTS as a potential incentive mechanism for accurate reporting
187 across a range of survey types (e.g. John *et al*, 2012; Weaver and Prelec, 2013). Incentives to
188 accurate reporting may be worth using where respondents believe (perhaps falsely) that there may
189 be real consequences induced by the findings of a survey and so there is the potential for the
190 hypothetical survey to be ‘incentive incompatible’ (i.e. for respondents to misreport their true
191 preferences; see Hurwicz, 1972 p.320), and also where reporting accurately is costly in terms of
192 time or cognitive effort or where respondents have little incentive to answer carefully and honestly.

193

194 The aim of the BTS approach is to elicit truthful (i.e. honest and carefully considered) subjective
195 data in situations where the objective truth is unknowable. The method uses an information
196 scoring system that is thought to elicit truthful answers from a sample of rational expected-value-
197 maximizing respondents (Prelec, 2004). It has been claimed that the approach can eliminate bias
198 common in contingent valuation studies when applied to responses pertaining to the respondent’s
199 contribution to a public good (Weaver and Prelec, 2013). The BTS can be described as follows.
200 People are asked a question (e.g. would you be prepared to pay an additional amount for
201 legislation x?). Additionally, they are asked to estimate what proportion of people (in general)
202 they think would give a particular response. Both these responses contribute to a formula made
203 known to the respondent, which integrates their responses with the responses of other individuals
204 asked the same questions. The BTS theorem contends that if individuals provide personally
205 truthful answers then expected scores are maximized. Respondents are then incentivised (usually

asked respondents whether they would be likely to be willing to pay something for the legislation on a five-point scale (Definitely yes, Possibly yes, Not sure, Probably not, Definitely not).

206 monetarily) to get a higher score. The precise nature of the formula is shown in the online
207 Appendix, available at the publisher’s website. The formula assigns high scores to respondents’
208 personal answers that are more common across the sample than collectively predicted by all
209 respondents when asked how they think others would respond (i.e. what proportion of people
210 generally would give a particular response). For example, an answer shared by 10% of all
211 respondents compared to a prediction of 5% would be surprisingly common and receive a high
212 score. The assumption is that people tend to believe that the proportion of responses that
213 correspond to their own response will be higher than the mean proportional responses elicited
214 from the whole sample. By rewarding people for giving ‘surprisingly common’ responses truth
215 telling is thus encouraged.

216
217 Half of the respondents were given a Bayesian Truth Serum Incentive (BTSI) statement to see if
218 it had an impact on their reporting (see Weaver and Prelec, 2013 for other examples of such
219 statements). The BTSI statement read: “Please note that one person will be chosen at random in a
220 prize draw to receive up to £100. The exact amount of the prize will be determined by the winner’s
221 Truth Score (the higher the score the higher the prize amount). The Truth Score is a measure
222 recently developed and published in the academic journal Science
223 (<http://nel.mit.edu/pdf/17BayesianTruthSerumcopy.pdf>). Even though only you know how
224 truthful your answers are, people who consider the questions carefully, answer honestly and take
225 care to avoid mistakes score higher on the Truth Score and provide more reliable information for
226 the survey.” For the BTS to work, respondents are not required to understand the mathematics of
227 the scoring nor the theory behind the BTS. However, they must believe that the method rewards
228 truth telling on average. If the BTSI had an impact, this should manifest itself as a difference in
229 the responses of incentivised vs non-incentivised individuals.

230

231 **4. Estimation and inference**

232 In addition to descriptive statistics, the principle method used within the study to estimate WTP
233 is a Bayesian Interval Regression model specified under two different assumptions concerning
234 the error distribution, i.e. the log-Normal and the Normal.

235

236 The two models are both of the standard linear form:

237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270

$$WTP_i = a + b BTSl_i + cZ_i + e_i \quad (1)$$

and

$$\ln(WTP_i) = a + b BTSl_i + cZ_i + e_i \quad (2)$$

where

BTSl_i = 1 if incentivised under BSTI and 0 otherwise
Z_i is a vector of covariates describing individual i
e_i is a normal error with mean 0 and variance σ².

The models were estimated with and without $b=0$ imposed, but we only present the full models below (in Table 4 which also outlines the covariates used) since all other coefficients were virtually the same with and without this restriction. The interval regression assumes that in either case the error is normally distributed with mean zero and constant variance. Interval regression is used to model dependent variables that have interval censoring. That is, each observation is known to lie within a range, but is potentially unobserved within that range. The upper and lower ranges can be bounded or unbounded, and if some of the values of some dependent variables are known exactly, then they can be treated as exact. An option to run interval regressions now exists for many packages including SAS, R, Gretl, and Stata. A full description of the interval model can be found in Stewart (1983). This model is the same whether it is estimated using a Classical or Bayesian approach. The former would map the likelihood function as expressed in Stewart (1983) and the latter can map the posterior (proportional to the likelihood multiplied by a weak prior) using any number of MCMC (Markov chain Monte Carlo) algorithms. The MCMC algorithms used here were coded from first principles in the GAUSS mathematical and statistical system software. Since WTP is only observed as an interval with an upper and lower bound (l,u), the WTP or $\ln WTP$ is treated as a latent variable with a conditional mean and variance as in the regressions above. Although the Bayesian model was used here with non-informative priors, the results will be virtually identical to those obtained using a Classical approach (Stewart, 1983). The log-Normal model is preferred in this case, given that the distribution of WTP estimates (as averages of the intervals) are clearly of a log-Normal type. However, as will be shown, the key conclusions of the analysis largely remain invariant to this model preference.

271 The impact of the BTSI on responses was examined in two ways. First, a test was applied to see
 272 whether the BTS scores of the incentivised group differed from the non-incentivised group.
 273 Then, the WTP distributions were examined to see if they were different across the two
 274 populations using an interval regression model with a BTSI dummy variable. The significance of
 275 this dummy variable would indicate whether incentivising participants had an impact on
 276 respondents' behaviour. In conducting this component normally distributed WTPs and log-
 277 Normal WTPs were allowed in order to see whether the results were robust to distributional
 278 assumptions.

279

280 **5. Results**

281 **Descriptive statistics**

282 Table 1 contains a summary of descriptive statistics for the sample of 665 responses. The sample
 283 contained 55% males and 45% females, with respondents living in households of 2.2 people on
 284 average, of whom 0.4 were under 16 years of age. Average household income was £33,500. Fifty-
 285 nine percent of respondents reported living in rural settings, i.e. in either villages or provincial
 286 towns, with 41% stating that they lived in urban areas. Nearly 16% of respondents reported being
 287 members of an organisation concerned with the welfare of animals such as the RSPCA or RSPB⁴.

288

289 Table 1 –Descriptive statistics for the survey sample

Description	Value(s)
Total sample size	665
Gender	Male (55%); Female (45%)
Age distribution (by group)	<40 (26.8%); 40-59 (32.6%) 60+ (40.6%)
Average household size (persons)	2.2
Average number of children in households <16 years	0.4
Average household income (£ p.a.)	33,500

⁴ This percentage would appear relatively high but it is likely that respondents interpreted this question quite widely including a range of organisations and relatively lax definitions of what constitutes membership. This percentage was adjusted downwards to 5% when accounting for sampling bias as described below.

Percentage of rural residents	59%
Percent of respondents who are members of an organisation concerned with the welfare of animals	16%
Percentage of respondents who were vegetarians	2.5%
Percentage of respondents who consumed chicken	94.9%
Consumption of meals (from all sources) containing chicken (number of meals / week)	1-2 (51.5%); 3-7 (43.2%); 7+ (5.3%)
Weekly spend on chicken (all sources) (£ /week / household)	0-£4.99 (37%); £5-£14.99 (54.5%); £15+ (8.5%)
Percent of sample making regular purchases of free-range chicken	22.9%

290

291 Almost 95% of respondents reported being consumers of chicken. Just under half of those who did
 292 not consume chicken stated their reason as being vegetarianism. Respondents consuming chicken
 293 reported eating an average of 3.05 meals per week containing chicken (mode 1-2; includes
 294 takeaways and meals out), with a range from one to 15. Respondents reported spending a modal
 295 value of £5-£10 per week (mean £7.8) on chicken purchases for their household, excluding
 296 takeaway meals and meals eaten outside the home (respondents were asked to select from
 297 categories for their response, which ranged from zero to more than £20 per week). Chicken from
 298 conventional production systems was by far the most common purchase, with free-range chicken
 299 regularly purchased by just 23% of the sample. Fifty-six percent of the sample reported purchasing
 300 free-range chicken meat only occasionally and 21% never purchased it.

301

302 Respondents were asked to rank their concern for farmed animals against a number of other widely
 303 held concerns, such as those in relation to the environment. Chicken eaters are most concerned
 304 about food safety, with animal welfare concerns near to last. Non-chicken eaters place food safety
 305 concerns last, but, again, animal welfare concerns are low in the order of priorities ('healthy diet'
 306 was ranked first). Respondents were asked to rank their level of concern for the welfare of broiler
 307 chickens reared in the UK on a scale of 0-10, where 0 = not concerned at all and 10 = very
 308 concerned. The sample average rank was 8.2, with chicken eaters scoring insignificantly higher

309 than non-chicken eaters, i.e. 8.6 compared to 8.2. There were significant age differences in ranking
310 of concern, with older respondents ranking concern more highly ($F=3.53$, $p=0.0147$).

311
312 **WTP elicitation**

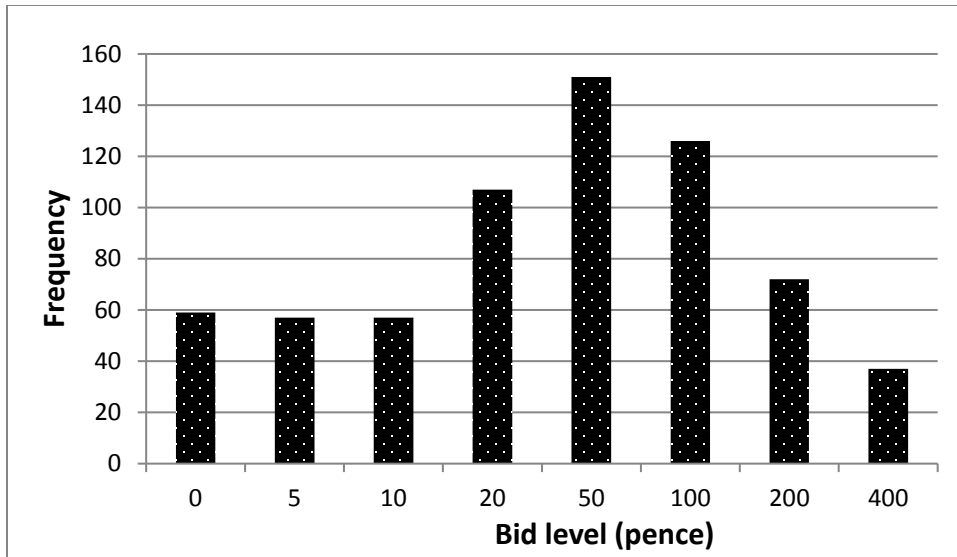
313 Respondents were asked to indicate the maximum sum that they would be willing to pay per
314 week as an additional amount on the cost of their chicken meat purchases (or, in the case of non-
315 chicken eaters, as an increase in taxation) to support the broiler legislation. Figure 1 shows the
316 distribution of WTP responses. Those classified as zero in Figure 1 are those who responded
317 'No' to the lowest WTP amount offered to them (5p per week extra) and may therefore include
318 those with a non-zero but low (i.e. less than 5p) WTP. In the estimation of WTP, in the case of
319 the unlogged data, the latent variables for those saying they would not pay the smallest possible
320 amount (5p) are truncated above by the smallest possible values in the survey (5p), and this
321 allows potentially negative WTPs. For the logged version, the logged WTP is truncated above
322 the log of the smallest possible value, except by construction the latent WTPs will (when being
323 anti-logged i.e. taking the exponential of the log WTP) imply positive though negligible WTPs.

324
325 Respondents were asked a number of debriefing questions to explore the reasons for their stated
326 WTP, and test the rationality of responses. These questions presented a number of propositional
327 statements and respondents had to express their level of agreement with each, using an 11-point
328 scale where 0 = 'does not reflect my views at all' and 10 = 'reflects my views completely' (see
329 Table 2). Respondents were first asked to indicate whether they understood the information
330 provided in the questionnaire about the broiler legislation. The mean sample rank score was 8.7,
331 indicating a high level of agreement. When asked whether they understood the WTP question (for
332 both payment vehicles), the great majority of respondents indicated that they did (mean score 8.9).

333

334 Figure 1. Distribution of sample over bid acceptance levels (pence).

335



336

337

338 Respondents generally believed that the welfare of chickens reared for meat needed to be improved
 339 and also expressed a belief that the Directive itself would act to improve the welfare of chickens
 340 reared for meat. It is interesting to note that although the Directive was introduced in 2010 (the
 341 study outlined here was part of an ex-post appraisal of the legislation) respondents were generally
 342 not aware of the Directive (for example, all focus group participants were found to be unaware of
 343 the legislation). There was a relatively high level of agreement by respondents that their WTP
 344 values reflected their concerns about the welfare of broiler chickens. Overall, respondents tended
 345 to agree with the statement that they should not have to pay more to improve chicken welfare. As
 346 might be expected, respondents who did not eat chicken were less likely to agree with the
 347 statement. There was no significant difference on this measure between those who regularly
 348 purchased free-range chicken and those who did not, even though there was a significant difference
 349 in WTP between these groups ($F=13.28$, $p<0.0001$), with regular purchasers willing to pay
 350 £0.9/week more than those who never purchased. Perhaps surprisingly, those respondents that
 351 most strongly expressed the view that they should not have to pay more (scores of 9 and 10) did
 352 not have a significantly lower WTP ($t=-0.04$, $p=0.9680$). Finally, most of the respondents
 353 (212/59% scored between 6 and 10) who received the questionnaire version with the BTSI agreed
 354 with the statement that they believed the information provided to them about the truth score (12%
 355 scored 5 suggesting they neither believed nor did not believe the information and 14% scored
 356 between 0 and 4 suggesting they did not believe the information).

357

358 Table 2. Responses to WTP debriefing questions, where average rank score expresses degree of
 359 agreement with propositional statements on a scale of zero to 10.
 360

Propositional statements	Rank score ^a (0–10)
I understood the information provided to me about the legislation	8.74
I understood the willingness to pay questions	8.94
The welfare of chickens reared for meat needs improving	8.36
The legislation will improve the welfare of chickens reared for meat	8.05
My WTP reflects my concern about chicken welfare	7.57
I should not have to pay more to improve chicken welfare	6.63
I believed the information provided about the Truth Score	6.23

361 ^a Note: A rank score of 0 = ‘does not reflect my views at all’ and a rank of 10 = ‘reflects my views completely’.
 362

363 Table 3 shows the numbers of respondents who said that they would or would not pay something
 364 for the legislation. It shows that although the majority (63.5%) say they would pay something,
 365 103 (15.5%) individuals reported that they were either definitely not or probably not prepared to
 366 pay. Of these, 58 subsequently went on to indicate a non-zero WTP amount when asked how
 367 much they would be willing to pay. This could reflect the WTP amounts presented to them,
 368 which started at £0.05 per week (0.6% increase on stated chicken expenditure). This may have
 369 induced a positive response from respondents who otherwise might have bid nothing. However,
 370 given this apparent inconsistency in stated preferences, for those respondents who had previously
 371 indicated that they would ‘definitely not’ or ‘probably not’ be willing to pay something for the
 372 legislation, a zero WTP is assumed in the estimation of WTP below to reduce any upward bias
 373 due to ‘yea saying’.

374

375 Table 3: Numbers of respondents who would be prepared to pay something for the broiler
 376 legislation

	Frequency	Percent
Definitely yes (1)	154	23.2

Possibly yes (2)	267	40.3
Not sure (3)	139	20.9
Probably Not (4)	77	11.6
Definitely Not (5)	26	3.9
Total	663	99.9

377

378

379 **Regression results**

380 Table 4 gives the interval regression results for both the Normal (in the bottom panel) and log-
381 Normal (in the top panel) models as described in Equations 1 and 2. Income, membership of an
382 organisation concerned with the welfare of animals (labelled RSPCA) and having children under
383 16 in the household all have a significant influence on WTP in both regressions. The regressors
384 have been demeaned so that for the Normal model the intercept can be interpreted as the mean
385 (and median since it is a symmetric distribution) estimate of WTP at the sample mean of the
386 descriptors. For the log-Normal model, the mean and median WTPs are calculated from the
387 intercept and variance of the log-Normal regression, but may still be interpreted as estimates at
388 the sample mean of the descriptors. The preferred model is the log-Normal, given that the
389 distribution of WTP estimates (as averages of the intervals) are clearly of a log-Normal type.

390

391 Table 4. Results from Normal and log-Normal interval regressions.^{a,b} Dependent variable is
392 WTP.

	Estimate	SD Est	Pseudo-t
<u>Log Normal</u>			
Intercept	3.33	0.05	65.17**
BTSI	0.08	0.10	0.86
ln(Income)	0.29	0.08	3.58**
ln(Age)	-0.01	0.14	-0.07
Female	0.19	0.10	1.82*
RSPCA	0.54	0.14	3.80**
Have Children	0.29	0.13	2.14**
Error Variance	1.65	0.10	
Median WTP	27.97	1.43	
Mean WTP	64.11	4.52	

<u>Normal</u>			
Intercept	56.43	3.07	18.34**
BTSI	4.95	6.13	0.80
ln(Income)	13.20	4.91	2.68**
ln(Age)	-0.71	8.50	-0.08
Female	5.17	6.47	0.80
RSPCA	30.51	8.55	3.56**
Have Children	21.37	8.36	2.55**
Error Variance	5782.59	372.53	
<hr/>			
Mean and			
Median WTP	56.43	3.07	

393 ^a Asterisks * and ** denote 2-tailed significance at the 10% and 5% significance respectively.

394 ^b Number of observations = 665

395

396 **WTP robustness checks**

397 In order to guard against ‘yea saying’, the models were re-estimated under the assumption that
398 those respondents who had previously indicated that they would ‘definitely not’ or ‘probably
399 not’ be willing to pay something for the legislation, had zero or even negative WTPs, even if
400 they subsequently indicated otherwise. In addition, we corrected for bias which may have been
401 caused by over-representation of some classes of socio-demographic stratification in the sample
402 compared with the general population. We did this by inputting average key socio-demographic
403 characteristics of the general population of England and Wales for 2014 (ONS, 2014a) into the
404 preferred model and re-estimating WTP. These key characteristics are income (£27,200), age
405 (40), gender (49.3% male), the proportion of households with children under 16 (30.3%) and an
406 assumption that 5% of the general population were members of an organization concerned with
407 the welfare of animals.

408

409 The preferred log-Normal model with these adjustments gives an estimated mean WTP of £62.5
410 and a median WTP of £21.5 per household per week. Differences between these values reflect
411 the highly diffuse upper tail in the log-Normal distribution. Thus, the median estimate is
412 considered the better measure of central tendency, as a measure of the consumer/citizen surplus

413 associated with the legislation⁵. These values compare with those of around £64 and £28
414 respectively of the unadjusted WTP estimates from the log-Normal model shown in Table 4.

415
416 However, to explore the robustness of this conclusion further the log-Normal model was re-
417 estimated assigning a zero WTP to 148 respondents who answered 'No' when asked (prior to
418 informing them about the legislation) if they would be willing to pay something to improve the
419 welfare of meat chickens but then subsequently stated a positive WTP to support the broiler
420 legislation. These results are not shown here but are very similar to the log-Normal model results
421 shown in Table 4 except that median WTP was estimated at £14.3 per household per week.

422

423 **Impact of the BTSI**

424 Before returning to the results in Table 4, it is worth noting that the mean BTS scores for
425 incentivised individuals relative to non-incentivised individuals were 0.17 and 0.13 respectively.
426 While the incentivised group had higher scores on average, a t-test of the difference between
427 group scores was not statistically significant (p -value > 0.15). With respect to the group of
428 individuals with a positive WTP, the incentivised and non-incentivised groups represented
429 69.5% and 71.2% of these respondents respectively. A z-test for differences in proportions was
430 again statistically insignificant (p value > 0.30).

431

432 Next we return to the results in Table 4 to test for the impact of the BTSI on WTPs. As before
433 two forms of interval regression were used, i.e. Normal and log-Normal (see Table 4). The point
434 here is not to make a comparison of the WTP results, but to show that inference is robust to
435 assumptions about the nature of the WTP distributions. As these are Bayesian estimates, the
436 "Pseudo-t" values are the ratio of the estimate divided by the standard deviation of the estimate,
437 similar to the t-statistic in classical regression. If this statistic exceeds 1.96 then it is likely that
438 the associated regressor is having a non-zero impact on the dependent variable. The BTSI
439 dummy indicates whether individuals have been incentivised according to the BTSI.

440

⁵ Note that the very few in the sample (34 of 665) who did not eat chicken had an average WTP of around twice that of chicken eaters.

441 The coefficient of BTSI can be interpreted as the increase in WTP (directly for the Normal or
442 logged WTP for the log-Normal case) resulting from the BTSI treatment. The coefficients are
443 positive for both the Normal (4.95) and log-Normal (0.09) cases, but given their large standard
444 deviations, these cannot be reliably said to have a positive impact on WTP. In addition to the
445 regression results shown in Table 4, we performed a regression for each distribution with the
446 BTSI dummy as the sole explanatory variable, in order to test that the lack of significance for
447 BTSI is not due to collinearity of the BTSI dummy with other regressors. Again, the coefficients
448 of BTSI were positive but not significant. From these results, it is evident that according to either
449 distribution there is no statistically-significant evidence that the use of the BTSI has had an
450 impact on WTP in this case.

451
452 In summary, the results show that the responses of people with respect to the truth score itself,
453 their WTP or the propensity to give inconsistent responses are largely invariant to the BTSI.

454

455 **6. Discussion**

456 Our survey results indicate that people in England and Wales generally support the broiler
457 Directive legislation and appear to have a substantial WTP to support it. WTP was found to be
458 positively correlated with income, level of concern about chicken welfare and belief that the
459 legislation will improve the welfare of meat chickens, as would be expected. Moreover, a large
460 majority of respondents stated that they understood both the information provided to them about
461 the Directive and the WTP questions they were asked to respond to. Thus, the survey instrument
462 appeared to work well and to result in credible responses from respondents. The sample size of
463 665 respondents meets the requirement recommended by Mitchell and Carson (1989 p. 229) for a
464 sample size of 600 usable responses or more for estimating benefits for policy purposes.

465

466 However, there are many opportunities for various types of bias within stated preference survey
467 instruments and there is a vast literature on such biases within stated preference research,
468 including the issue of misreporting. In the first place, since the legislation already exists, fully
469 informed and economically rational participants would not be expected to be willing to pay
470 anything more for what they already have and are already paying for, one way or another.

471 However, they may be willing to pay an additional amount to retain the legislation rather than it

472 being removed (i.e. up to the point where their consumer surplus is zero). Harvey and Hubbard
473 (2013) provide a critique of the elicitation of people's WTP in relation to animal welfare together
474 with a more general thought-provoking consideration of the application of economics to animal
475 welfare. Emotive issues such as animal welfare are susceptible to social desirability bias, where
476 respondents feel pressured to provide a socially desirable response (i.e. to show concern for the
477 humane treatment of animals), which may have been provoked by some questions in our survey.
478 We also acknowledge the possibility of starting point bias and/or range bias associated with the
479 WTP values presented to participants (where they may feel that either the first amount or the
480 range of amounts presented to them is some indication of the amount they should be willing to
481 pay). In addition, 'strategic bias' may have been present, especially in relation to the scenario
482 presented to respondents which described the legislation as being currently evaluated (which was
483 true). Respondents may have felt that implementation of the legislation was under threat and may
484 therefore have strategically over-stated their WTP to ensure that the legislation was
485 implemented. We have used a relatively cautious approach to WTP estimation to guard against
486 these potential biases.

487
488 Otherwise, responses of individuals may be shaped (perhaps resulting in misreporting) by the way
489 in which questions are presented (Kahneman and Tversky, 1984), the information they are
490 presented with (Ajzen *et al.*, 1996; Hensher, 2006), hypothetical bias (the difference between true
491 WTP and stated WTP) including the perceived social desirability associated with questions
492 (Nederhof, 1985) such as humane treatment of chickens in this survey, the time they are given to
493 complete the survey (Conte *et al.*, 2016), the length of the survey (Savage and Waldman, 2008)
494 and so on. It is not possible to test for all possible biases, priming or other effects of survey design.
495 Our survey instrument was carefully designed, pre-tested and piloted to minimize bias in responses
496 and the results analysed and presented to avoid estimation bias. Responses were scrutinised in
497 relation to their credibility, rationality and consistency to identify potential issues that might cast
498 doubt on the validity of the survey instrument design, people's responses to the questions and the
499 subsequent results. Questions regarding respondents' consumption of chicken and their attitudes
500 toward animal and chicken welfare were placed prior to information on the Directive and
501 elicitation of WTP to encourage respondents to think about the importance of chicken in their diets

502 and the importance that they give to animal and chicken welfare (e.g. in relation to other concerns
503 such as food safety and a healthy diet).

504
505 Loomis (2014) identifies and reviews both *ex ante* and *ex post* strategies for overcoming
506 hypothetical bias in stated preference surveys. The former includes ‘consequentiality designs’
507 which means that the survey has some potential effect on the utility of respondents such as higher
508 prices or taxes. This approach was used in the survey reported here together with binary,
509 dichotomous choice question formats and compulsory payment mechanisms as recommended by
510 Loomis (see also Carson and Groves, 2007). ‘Honesty and realism’ approaches are also
511 recommended such as the one used in this survey which involved a request to respondents to
512 “consider your answers to questions carefully, answer honestly and take care to avoid mistakes”.
513 In addition, the BTSI was used to further encourage truthful responses. To reduce the possibility
514 of hypothetical bias respondents were also reminded of their budget constraints. Loomis suggests
515 that social desirability bias and cognitive dissonance (where respondents gain utility by responding
516 according to perceived social norms rather than their own personal values) are reduced by having
517 multiple bid values (with small positive amounts) as used in this survey, and by making responses
518 impersonal and anonymous (again as used in the internet survey of this study). In addition, asking
519 respondents what they think others would pay is a way of potentially gauging over-statement of
520 WTP due to social desirability bias (see Lusk and Norwood, 2009) – an approach also used in this
521 study and used as an integral part of the BTS. Ex post methods to reduce hypothetical bias
522 identified by Loomis include reporting median WTP to minimize the effect of implausibly high
523 WTPs, recording respondent uncertainty in relation to their WTP responses and recoding of ‘Yes’
524 responses as ‘No’ where there is sufficient uncertainty regarding respondents having a true positive
525 WTP, all of which were used in this study. It is also noted that using a private good (such as chicken
526 meat in this study) reduces hypothetical bias. Loomis warns that there is no consensus regarding
527 the best method to correct for hypothetical bias, that measuring hypothetical bias is difficult (the
528 analyst needs to know the true WTP) and that it is possible to over-correct for this bias and so
529 underestimate WTP.

530
531 We estimate the median WTP as a measure of the consumer surplus associated with the broiler
532 legislation, as £21.5/household/year from the preferred log-Normal Bayesian Interval Regression

533 model, after correcting for bias, representing 5% of the total annual amount that respondents
534 estimated they spent on chicken meat. Our more conservative estimate, assuming a zero WTP for
535 those who said they would not be willing to pay more, despite subsequently providing a positive
536 response, is £14.3 per household per week.

537
538 According to the UK National Census (ONS, 2014a), there were 23.4 million households in
539 England and Wales in 2011. Using the WTP estimate of £21.5/year from above, the aggregate
540 amount that consumers in England and Wales would be willing to pay for the broiler Directive is
541 £504 million/year. The more conservative estimate generates aggregate benefits of £333 million/
542 year. A 2014 survey of 119 commercial conventional broiler producers estimated the cost of
543 compliance with the broiler Directive in England and Wales over the period 2010 to 2014 to be
544 £108.4 million (current prices), an average of approximately £21.7 million per year or nearly £21.9
545 million per year including government inspection and enforcement costs (Defra, 2017; Defra
546 2010), implying a benefit-cost ratio of 23:1. The more conservative estimate gives a benefit-cost
547 ratio of 15:1. Some radical changes in assumptions for both benefits and costs would be required
548 to conclude that the costs of the legislation exceed the benefits.

549
550 Prior to the introduction of the Directive, SAC Commercial Ltd (SAC, 2005) estimated that
551 consumers in England and Wales would be willing to pay the equivalent of just under
552 £20/household/year (2014 prices) which is similar to the estimate generated above.⁶ However,
553 people's perceptions of the value of the legislation may have been somewhat different before and
554 after implementation of the legislation so the two estimates are not wholly comparable.

555
556 With regard to the BTS, we found no statistically significant impact of the BTSI on responses to
557 the survey. We believe this negative finding is at least as important as those studies which
558 suggest that the BTS has an impact. While we would encourage further investigation of this
559 approach, we believe that the literature should report cases where this technique has not
560 demonstrably changed behaviour in addition to occasions when it has. Not to do so would lead to

⁶ The SAC estimate of £7.53/person/year. According to the National Census (ONS, 2014a), average household size in 2013 was 2.37 persons with 1.86 adults per household. Therefore, the SAC (2005) WTP estimate equates to £14/household/year, in 2005 prices. Adjusting for inflation (ONS, 2014b), the SAC estimate is just under £20/household/year.

561 a biased representation of the weight of evidence in favour of positive BTS effects (more
562 generally, for the importance of reporting negative findings we refer readers to the debate
563 concerning ‘priming’ (Shanks *et al.* 2013)).

564
565 There are a number of reasons why our study may have failed to observe a significant BTSI
566 impact. First, while there may have been misreporting by the sample, the level or nature of the
567 incentive may not have been sufficient to alter respondent behaviour. More specifically, the
568 incentive provided for truth telling (the ‘prize draw’) may have been too small and uncertain,
569 with respondents perceiving only a small chance of gaining financially for ‘truth telling’ (with an
570 unknown probability of winning because respondents did not know the number of people taking
571 part in the prize draw). Second, although most respondents had some level of belief in the truth
572 score there was an element of disbelief which may have reduced the impact of the BTSI across
573 the sample. Third, irrespective of the efficacy of the incentive, there may simply have been no
574 significant misreporting in the sample and therefore no error to correct. This effect may be the
575 case because the perceived gravity of the issue being investigated encourages truthful reporting.
576 It is worth noting that some previous studies which have found the BTS to be effective asked
577 respondents to report on what might be deemed more trivial matters, for example surveys of the
578 extent of people’s recognition of world leaders’ names and film titles (see Weaver and Prelec,
579 2013). Arguably, the more gravity that respondents attach to the issue under investigation, the
580 more likely respondents are to carefully consider the questions presented to them, thus reducing
581 one potential source of misreporting. Additionally, a number of aspects of the survey and
582 questionnaire were designed to minimise bias and reduce any potential misreporting and this
583 may have contributed to an absence of misreporting by respondents.

584 585 **7. Conclusion**

586 We find that people in England and Wales have a substantial stated WTP to support legislation to
587 monitor and improve the welfare of broiler chickens in the UK. At a societal level, the benefits
588 of the legislation, as measured by people’s estimated WTP, greatly outweighed the costs (as
589 estimated by a survey of broiler producers). This suggests both that the current broiler legislation
590 was worth implementing and that it is worth continuing in the UK. Use of a Bayesian Truth
591 Serum Incentive (BTSI) method to encourage honest and careful responses made no statistically

592 significant difference in terms of people's responses or their WTP. This finding suggests that
593 further studies need to be undertaken to better determine the exact circumstances under which
594 the use of the BTSI has most effect in reducing misreporting.

595

596 **Supporting Information**

597 Additional Supporting Information may be found in the online version of this article:

598 **Appendix S1** Specimen copy of the online questionnaire

599 **Appendix S2** Specification of the Bayesian Truth Serum scoring formula

600

601 **References**

602 Ajzen, I., Brown, T.C. and Rosenthal, L.H. 'Information bias in contingent valuation: effects of
603 personal relevance, quality of information and motivational orientation', *Journal of*
604 *Environmental Economics and Management*, Vol. 30, (1996) pp. 43-57.

605

606 Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M.,
607 Loomes, G., Mourato, S., Ozedemiroglu, E., Pearce, D., Sugden, J. and Swanson, J. *Economic*
608 *valuation with stated preference techniques: A manual* (Cheltenham, UK: Edward Elgar, 2002).

609

610 Bennett, R., Kehlbacher, A. and Balcombe, K. 'A method for the economic valuation of animal
611 welfare benefits using a single welfare score', *Animal Welfare*, Vol. 21(S1), (2012) pp. 125-130.

612

613 Bennett, R. and Appleby, M. *Animal welfare policy in the European Union*. In: Oskam, A.,
614 Meester, G. and Silvis, H. (eds.) *EU policy for agriculture, food and rural areas*. 2nd edition.
615 Wageningen Academic Publishers, (2011) pp. 249-258.

616

617 Bennett, R., Butterworth, A., Jones, P., Kehlbacher, A. and Tranter, R. *Valuation of animal*
618 *welfare benefits*. Report to the Department for the Environment, Food and Rural Affairs
619 (University of Reading: Reading, UK, 2011).

620

621 Carson, R.T. and Groves, T. 'Incentive and informational properties of preference questions',
622 *Environmental and Resource Economics*, Vol. 37(1), (2007) pp. 181-210.
623

624 Clark, B., Stewart, G.B., Panzone, L.A., Kyriazakis, I. and Frewer, L.J. 'A systematic review of
625 public attitudes, perceptions and behaviours towards production diseases associated with farm
626 animal welfare', *Journal of Agricultural and Environmental Ethics*, Vol. 29(3), (2016) pp. 455-
627 478.
628

629 Conte, A., Scarsini, M. and Sürücü, O. 'The impact of time limitation: insights from a queueing
630 experiment', *Judgment and Decision Making*, Vol. 11(3), (2016) pp. 260-274.
631

632 Defra. *Study to evaluate the effectiveness of Regulation (Directive 2007/43/EC) in England and*
633 *Wales*. EVID4 Evidence Project Final Report (Rev. 06/11) (2017).
634

635 Defra Council Directive 2007/43/EC. *Laying down minimum rules for the production of chickens*
636 *kept for meat*, Impact Assessment No. Defra1000 (Defra: London, 2010).
637

638 European Commission Council Directive 2007/43/EC. *Laying down minimum rules for the*
639 *protection of chickens kept for meat production*. Official Journal of the European Union, L 182/19,
640 12.7.2007. (2007).
641 <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32007L0043&rid=4>
642

643 European Commission, Animal Welfare. (2016)
644 http://ec.europa.eu/food/animals/welfare/index_en.htm
645

646 Harvey, D. and Hubbard, C. 'Reconsidering the political economy of farm animal welfare: An
647 anatomy of market failure', *Food Policy*, Vol. 38, (2013) pp. 105-114.
648

649 Heinzen, R.R and Bridges, J.F. 'Comparison of four contingent valuation methods to estimate
650 the economic value of a pneumococcal vaccine in Bangladesh', *International Journal of*
651 *Technology Assessment in Health Care*, Vol. 24(4), (2008) pp. 481-487.

652
653 Hensher, D.A. ‘How do respondents process stated choice experiments? Attribute consideration
654 under varying information load’, *Journal of Applied Econometrics*, Vol. 21, (2006) pp. 861-878.
655
656 Hurwicz, L. ‘The design of mechanisms for resource allocation’, *American Economic Review*,
657 Vol. 63, (1972) pp. 1-30.
658
659 John, L.K., Loewenstein, G. and Prelec, D. ‘Measuring the prevalence of questionable research
660 practices with incentives for truth telling’, *Psychological Science*, Vol. 23(5), (2012) pp. 524-
661 532.
662
663 Kahneman, D. and Tversky, A. ‘Choices, values, and frames’, *American Psychologist*, Vol. 39,
664 (1984) pp. 341–350.
665
666 Lagerkvist, C.J. and Hess, S. ‘A meta-analysis of consumer willingness to pay for farm animal
667 welfare’, *European Review of Agricultural Economics*, Vol. 38(1), (2011) pp. 55-78.
668
669 Loomis, J.B. ‘Strategies for overcoming hypothetical bias in stated preference surveys’, *Journal*
670 *of Agricultural and Resource Economics* Vol 39(1), (2014) pp. 34-46.
671
672 Lusk, J.L. and Norwood, F.B. ‘An inferred valuation method’, *Land Economics* Vol 85, (2009)
673 pp. 500-514.
674
675 Mayfield, L.E., Bennett, R.M., Tranter, R.B. and Wooldridge, M. J. ‘Consumption of welfare-
676 friendly food products in Great Britain, Italy and Sweden, and how it may be influenced by
677 consumer attitudes to, and behaviour towards, animal welfare attributes’, *International Journal*
678 *of Sociology of Food and Agriculture, Food and Agriculture*, Vol. 15(3), (2007) pp. 59-73.
679
680 McFadden, D. Contingent valuation and social choice. *American Journal of Agricultural*
681 *Economics* Vol 76, (1994) pp. 689-708.

682 Mitchell, R.C. and Carson, R.T. *Using surveys to value public goods. The Contingent Valuation*
683 *Method.* (Resources for the Future: Washington D.C., 1989)

684 Nederhof, A.J. ‘Methods of coping with social desirability bias: a review’, *European Journal of*
685 *Social Psychology*, Vol 15, (1985) pp. 263-280.

686

687 ONS, *Families and Households, 2013.* (2014a)
688 <http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-328237>
689

690 ONS, *Consumer Price Inflation*, September 2014. (2014b)
691 <http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-323601>
692

693 Pearce, D. and Ozdemiroglu, E. *Economic valuation with stated preference techniques*,
694 (Department for Transport, Local Government and the Regions: Rotherham, 2002).

695

696 Prelec, D. ‘A Bayesian truth serum for subjective data’, *Science*, Vol. 306, (2004) pp. 462-466.
697

698 SAC, *Estimating non-market benefits of reduced stocking density and other welfare increasing*
699 *measures for meat chickens in England*, Report to Department of Food and Rural Affairs: Project
700 Code AW0236, 2005.

701

702 Savage, S.J. and Waldman, D.M. ‘Learning and fatigue during choice experiments: a comparison
703 of online and mail survey modes’, *Journal of Applied Econometrics*, Vol. 23(3), (2008) pp. 351-
704 371.

705 Shanks, D.R., Newell, B.R., Lee, E.H., Balakrishnan, D., Ekelund, L., Cenac, Z., Kavvadia, F.
706 and Moore, C. ‘Priming intelligent behavior: an elusive phenomenon’, *PLoS ONE*, Vol 8(4),
707 (2013), e56515.

708

709 Stewart, M. B. On least squares estimation when the dependent variable is grouped. *Review of*
710 *Economic Studies* Vol 50, (1983) pp. 737-753.

711

712 Weaver, R. and Prelec, D. ‘Creating truth-telling incentives with the Bayesian Truth Serum’,
 713 *Journal of Marketing Research*, Vol. 50(3), (2013) pp. 289-302.

714
 715 Zhao, J and Kling, C.L. Willingness to pay, compensating variation and the cost of commitment.
 716 *Economic Inquiry* Vol 42(3), (2004) pp. 503-517.

717
 718
 719
 720

721 **Appendix 2: BTS (Bayesian Truth Serum) scoring formula**

722
 723 If question answers and predictions are denoted by

724
 725 $x^r = (x_1^r, \dots, x_m^r)(x_k^r \in \{0, 1\}, \sum_k x_k^r = 1)$ (1) and

726
 727 $y^r = (y_1^r, \dots, y_m^r)(y_k^r \geq 0, \sum_k y_k^r = 1)$ (2) respectively,

728
 729 where x^r are the answers of respondent, r to questions 1 to m and y^r are the predictions of
 730 respondent, r to questions 1 to m , then we can calculate the population endorsement frequencies,
 731 \bar{x}_k , and the (geometric) average, \bar{y}_k , of predicted frequencies as:

732
 733
$$\bar{x}_k = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{r=1}^n x_k^r, \quad (3)$$

734
 735
$$\log \bar{y}_k = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{r=1}^n \log y_k^r \quad (4)$$

736
 737 where n is the sample size. Instead of applying a preset answer key, we evaluate answers
 738 according to their information score, which is the log-ratio of actual-to-predicted endorsement
 739 frequencies. The information score for answer k is

740
 741
$$\log \frac{\bar{x}_k}{\bar{y}_k} \quad (5)$$

742
 743 The total score for a respondent combines the information score with a separate score for the
 744 accuracy of predictions:

745
 746 score for respondent r = information score + prediction score =

747
 748
$$\sum_k x_k^r \log \frac{\bar{x}_k}{\bar{y}_k} + \alpha \sum_k \bar{x}_k \log \frac{y_k^r}{\bar{x}_k}, 0 < \alpha$$
 (6)

749
750
751 (Prelec, 2004).
752
753
754