

## **Annex 16**

### **Provision of a method for the economic valuation of animal welfare benefits suitable for use in policy appraisal**



**Extended Detailed Report**

**School of Agriculture, Policy and Development  
Department of Agri-Food Economics and Marketing**

**January 2025**

<b>Contents</b>	<b>Pages</b>
1. Introduction – background, policy need, aims and objectives	2 - 3
2. Welfare assessment: approach, methods and results	3-10
3. Economic valuation: approach, methods and results	10-26
4. Use of results for economic valuation of the benefits of policy scenarios	26-31
5. Discussion: robustness of the method and of results, limitations	31-33
6. Conclusion	33-34
7. References	34-35

## **Annexes**

1. Description of the IDEA protocol (Hemming et al, 2018)
2. Guidance provided to the expert panel
3. Policy case study specifications
4. Choquet scores elicited from the expert panel
5. Results of the expert panel scoring of the policy case studies
6. (i) Individual scores and comments of expert panel members  
(ii) Summary of expert panel rationale for policy scenario scores
7. The first exploratory survey questionnaire (CV)
8. The second exploratory survey questionnaire (CE1)
9. The third exploratory survey questionnaire (CE2)
10. The fourth exploratory survey questionnaire (CE3)
11. Questionnaires and transcripts of focus group discussions
12. Transcripts of verbal protocol interviews
13. Example copy of the main survey questionnaire
14. Main survey quotas and their achievement
15. Wtp Valuation Tables

## **Introduction**

### **Background and policy need**

There is a requirement to assess the costs and benefits associated with government policy and policy changes (HM Treasury Green Book, 2025). This is true for policy that impacts on animal welfare as well as for other areas of government policy. Although there is a generally accepted way of assessing the costs of policy, there is no single accepted or standardized method for valuing the benefits of policy designed to improve the welfare of animals. Moreover, there is no method where the economic valuation estimates (in relation to the same species/animal type) are transferrable between policies and policy appraisals. This means that every different assessment of changes to animal welfare brought about by an intervention or change in policy requires a new benefit valuation study from which to derive benefit estimates. This is time consuming, costly and provides significant inconsistency between policy appraisals making comparison of alternative policies more difficult.

Previous Defra-funded research (Bennett et al, 2012; Kehlbacher et al 2012, 2013) explored a stated preference approach for the valuation of animal welfare benefits using a single welfare score on a 0-100 scale. Choice experiment and contingent valuation methods were used to elicit citizens' willingness to pay for increases in the welfare score for three different farm animal species. However, this research was a relatively small pilot study of 300 citizens, considered only three types of livestock product and did not develop the welfare assessment method needed to generate the welfare scores associated with different policies.

### **Aim and objectives of the project**

The overall aim of the project described in this report was to further develop the work above to provide a protocol for the economic valuation of the benefits associated with improvements to the welfare of farm animals brought about by changes in policy affecting animal production from birth to slaughter. The outputs from using this economic valuation method could then be used as inputs to policy appraisal to support policy decisions.

Specifically the objectives of the project were to:

1. Develop a welfare assessment protocol to enable the scoring of changes in the welfare of farm animals as a result of changes in the way in which animals are used and kept.
2. Provide a robust valuation method for valuing changes in animal welfare from the scoring protocol above, using stated preference techniques to elicit consumer and citizens' willingness to pay.
3. Undertake a national study to derive welfare benefit estimates that are transferable across policy areas and that can be used in policy appraisal.
4. Apply the policy tool to the appraisal of six case-study policy changes (as agreed with Defra).

## **Work streams**

The project consisted of two main work streams.

The first work stream was concerned with the development, refinement and testing of a method for the assessment of changes to farm animal welfare as a result of a policy change. This involved the development and implementation of a protocol for an expert panel of animal welfare scientists to provide welfare assessment scores for specified animal production scenarios on a 0-100 scale, where zero denotes extreme suffering for the animal and 100 denotes the highest achievable welfare possible. In this way, potential improvements to animal welfare, for example, brought about by changes in policy and animal production methods, could be assessed in relation to current practice.

The second work stream involved economic valuation of animal welfare impacts by using choice experiment surveys to elicit people's willingness to pay to increase the welfare score of animals in the UK. This further developed the methodology presented in Bennett et al (2012) described above. Following testing of the economic valuation method, a final valuation protocol was used to generate estimates for a one point increment in the animal welfare score for each of the six main farmed species/production systems (broilers, laying hens, indoor pigs, dairy cows, beef cattle and sheep). The welfare score provides an important common link between the welfare assessment and the economic valuation parts of the project to enable economic valuation of improvements to animal welfare brought about by changes in policy.

It should be noted that the welfare assessment method developed for this project is for use in relation to farm animals and would not be directly transferable to other types of animals such as companion animals, laboratory animals, zoo animals or wild animals. However, it may be possible for protocols using the same general methodology to be developed for these applications.

## **Welfare Assessment**

This section first describes the method used for assessment of the animal welfare impacts of policy in terms of the welfare score. It then presents some results from using the method in relation to six policy case studies.

### **Approach and Methods**

The welfare assessment work stream involved the completion of nine tasks as outlined below.

1. Description of the welfare assessment task and the work of the expert panel.
2. Documentation of the welfare assessment protocol with instructions and guidelines for expert panel members.
3. Recruitment of independent expert panel members.
4. Scheduling of welfare assessment scoring tasks in relation to six policy case studies.
5. Induction of panel members in the work of the panel and the welfare assessment protocol.

6. Production of specifications for the six policy case studies and the welfare assessment scenarios contained within them.

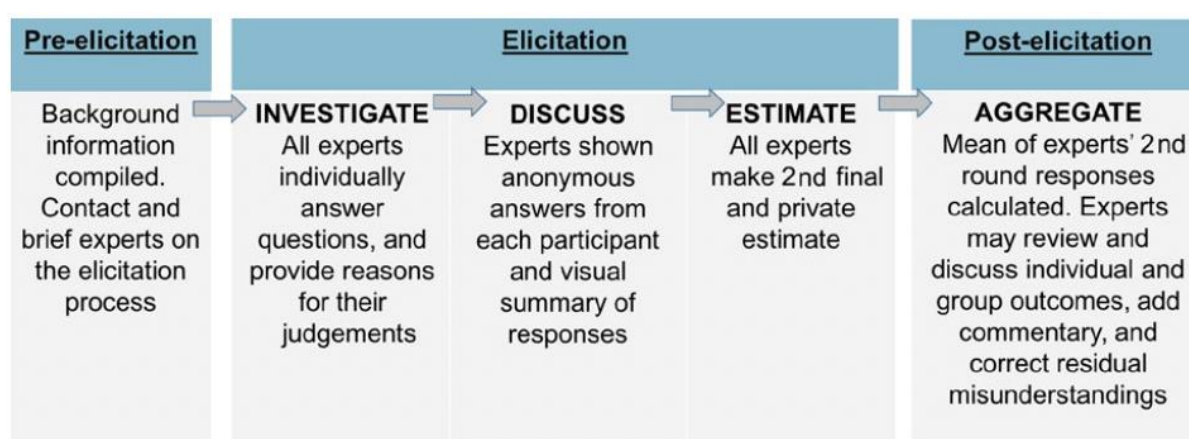
7. Production and sharing of information with the expert panel in relation to each of the policy case studies and scenarios.

8. Scoring of each policy scenario by the expert panel and collation of scores (see below).

9. Review of welfare assessment scores by the panel.

The welfare assessment work of the expert panel was designed according to the IDEA structured elicitation protocol which has been widely used and well-documented (Figure 1), for example in the area of environmental assessment using expert panels (Hemming et al, 2018). Annex 1 provides a more detailed description of the IDEA protocol taken from Hemming et al (2018).

**Figure 1. The IDEA protocol adapted from Burgman (2015)**



Hemming et al (2018) suggest that 12 experts are sufficient for an expert panel with little benefit from recruiting more. Thirteen independent animal welfare science experts were recruited to the welfare assessment panel for the project with a minimum of two specialists for each species/system, to allow for times when a panel member might be unavailable. The species/systems considered were laying hens, broiler meat chickens, indoor pigs, sheep, dairy cattle and beef cattle.

Panel members attended a half-day online induction session. During the session, the assessment protocol was presented and discussed, panel members undertook a practice scoring session followed by discussion of the protocol and suggested improvements to it. The latter became an ongoing discussion with and between panel members so that the protocol was continually being reviewed and where appropriate, improvements were made over the course of the project. These improvements did not materially change the fundamental protocol allowing scores generated for different policy scenarios over the course of the project to be compared.

The detail of the welfare assessment protocol is described below.

The welfare assessment protocol uses a framework which draws on the Five Domains of Animal Welfare model together with the welfare Criteria and Principles of the Welfare Quality protocols.

The Five Domains model (nutrition, environment, health, behaviour and mental state) is a widely accepted framework for considering animal welfare and for its assessment (see Mellor, 2017). In addition, in the context of farm animal welfare, the Welfare Quality (WQ) protocols have been widely used and adapted to facilitate on-farm welfare assessment. Welfare Quality developed from a large EU-funded project involving multiple scientists across Europe and other countries. It developed an on-farm welfare assessment tool with a series of welfare protocols for different farm production systems (see <http://www.welfarequality.net>). The WQ protocols score welfare according to 12 Criteria which are then grouped into four Principles (which are very similar to the first four of the domains in the Five Domains model). The Criteria and Principles are the same for every species/system considered. However, the WQ protocols are designed to be used for on-farm assessment, where an assessor visits a farm and uses a number of specific measures to determine welfare scores for each of the Criteria. WQ uses two or three on-farm measures to determine each Criteria score such as the number of working drinkers that animals have access to, the space that animals have, the injuries they have or the prevalence of particular diseases and conditions (e.g. lameness).

The welfare assessment protocol developed for this project differs to that used by WQ. Although the Criteria and Principles look similar in terms of their particular welfare focus (e.g. on how well animals are fed or watered, how much space they have, injuries and disease, ability to display normal behaviours) the Criteria descriptions are much expanded to include all aspects of the Criteria that impact on welfare not only the limited measures used in the WQ protocols. This includes consideration of the affective states of animals and opportunities for positive welfare experiences. This effectively integrates the fifth domain of the Five Domains model into consideration of each of the welfare Criteria and into the welfare assessment protocol. The protocol also includes the whole life experience of animals from birth to slaughter (unlike WQ which only considers on-farm welfare). An example score sheet, which is populated with the first and second-round mean panel member scores for colony cage egg production is shown in Table 1 below. The table shows descriptions of each of the welfare Criteria together with the welfare Principles. Guidance issued to the panel as to how they should approach the scoring exercises is shown in Annex 2.

Panel members are sent (by the project facilitators) by email the specification of a policy case study together with information about the production systems involved two to three weeks prior to an online panel meeting to discuss their scoring (see Annex 3), together with instructions and reminders regarding the assessment protocol. Panel members are able to share information between each other (a project Sharepoint site was set up to facilitate this) but were instructed not to discuss their scores. Panel members could seek clarification from the project facilitators regarding specification of the policy case study and its scenarios. For example, they might seek clarification on the boundaries of the systems being considered in terms of types of animals or farms, types of husbandry practices etc. All of the policy case studies considered all animals involved in the specified scenarios across the UK, including the range of farms and practices found across the nation.

Panel members are asked to provide 'first round' scores for each of the welfare Criteria in relation to each of the policy scenarios being considered for that particular policy case study. They also supply comments on their scoring for each of the Criteria (i.e. the rationale behind their score and particular considerations that they felt important to the score). They then return their scores to the facilitators and the scores are collated and fed back to the panel so that members are able to see

each other's anonymized scores prior to meeting to discuss. Panel members are also provided with aggregated panel scores for each Criteria, each of the four Principles and as a single score from combining the four Principle scores. The aggregate values of Principles and the single score were shown in two ways: as a simple arithmetic mean and as a weighted average using a Choquet Integral function which contains the panel's score weightings (see below for further details regarding this Choquet Integral function).

**Table 1 Example expert panel welfare assessment score sheet with a panel members' first and second round scores for colony cage egg production in the UK**

Scenario 1 - Colony cage egg production					
Score sheet					Name of scorer: Expert 11
Welfare principles	Welfare criteria		1st round score	2nd round score	Comments Round 1
	<div><div>0 – Lowest level of welfare</div><div>Highest level of welfare = 100</div><div></div></div>				
Good feeding	1	Provision and access to food. Animals should have appropriate access to the quantity and quality of appropriate foodstuffs for health and wellbeing.	55	60	Birds given diet appropriate for good health but not for pleasure or choice such as foraged food on the range, natural ingredients such as insects, or preventative medicinal foodstuffs they could self-medicate with for good health e.g. herbs.
	2	Provision and access to water. Animals should have appropriate access to the quantity and quality of water for health and wellbeing.	70	70	Birds have continuous supply of water but access might be hindered due to lack of space, also they lack choice of water sources, personal observation suggests birds like drinking from natural sources on
Good housing	3	Animals should have comfort when resting.	40	40	Birds are only able to perch, or to lie on wire, no choice of substrates or perches to rest on, restricted room to stretch out/find comfortable positions.
	4	Animals should have thermal comfort being neither too hot nor too cold.	60	60	Units kept at even temperatures which are neither too hot nor too cold, but choice of comfortable temperatures not given.
	5	Animals should have sufficient space to move freely.	20	20	Birds given more space in enriched cages to move, but not freely.
Good health	6	Animals should be free from injuries and disorders (e.g. skin conditions, lameness, bone fractures etc.).	40	45	Restricted movement may mean reduced injuries, but feather pecking remains an
	7	Animals should be free from disease, including metabolic conditions, with high standards of health care and hygiene.	60	50	Possible to have good hygiene standards in cages.
	8	Animals should not suffer pain - for example as a result of poor management, handling, surgical or other procedures, slaughter etc.	50	50	Behavioural restriction may mean reduced opportunity for pain, nothing about the system suggests making poor handling, surgery, slaughter more likely.
Appropriate behaviour	9	Animals should be able to express normal, non-harmful social behaviours (such as grooming and social bonding).	20	20	Close confinement makes harmful social interactions more likely.
	10	Animals should be able to express other normal behaviours (e.g. foraging, exploring).	10	15	Very little opportunity to express these behaviours due to lack of space and resources especially choice of resources.
	11	Animals should be handled well with positive and not negative animal-human relationships.	10	10	Very little/no opportunity to foster good human-animal relationships.
	12	Additional aspects not already adequately covered above in relation to the balance between positive and negative affective states for animals.	10	10	Negative emotions such as frustration, boredom apathy more likely, very little/no opportunity for promoting positive

The panel then meets online (usually a few days after receiving the collated scores) to discuss the scores. The panel could have met in-person but the online method was much easier to schedule and much less costly. This meeting is chaired by a facilitator supported by the project lead and a person responsible for collating the scores. The facilitator leads discussion regarding the panel's scores going through each of the Criteria one by one. Prior to this, panel members are reminded that they will have the opportunity to revise their scores if they wish, in the light of the discussion, in a second round of scoring and so should take notes as to how they might want to change their score as each Criteria and its score is discussed. There was no requirement for panel members to reach consensus as to their scores. Each member was free to decide their own scores.

Panel members are then asked to provide their second-round scores during the meeting by submitting new spreadsheets with their scores to the person collating the scores. The panel was given around 20-30 minutes to do this depending on the nature (e.g. complexity) of the policy scenarios needing to be scored. Collated and aggregated second-round scores are reviewed and discussed by the panel in the meeting. The panel considers whether they feel that the final aggregate scores adequately represent their joint perspectives regarding the welfare level of animals in the scenarios considered.

All six case-study online meetings with the panel were recorded and transcribed to allow later reference including reference to them to produce summaries of the main welfare considerations discussed for each Criteria.

### **Choquet Integral weightings**

The Choquet Integral aggregation function was used to enable different Criteria scores to be combined, taking into account how the welfare of animals might be impacted by relatively high or low scores in particular Criteria. The reasoning behind this is unweighted arithmetic means fail to take sufficient account of the degree to which a high welfare score in one Criteria can or cannot compensate for a relatively low score in another when computing combined scores of the Criteria. In order to estimate the weightings to be used, combined scores were elicited from the panel for the combinations of Criteria scores that determined each of the Principle scores. For efficiency of elicitation of combined scores from the panel and the estimation of the resultant weightings, this was done by using scores of 70/100 and 30/100 for each Criteria combination.

For example, the Principle of 'Good Feeding' is determined by the scores of Criteria 1 and 2 which are 'provision and access to food' and 'provision and access to water'. Panel members were individually asked, "if the Criteria score was 70 for Criteria 1 and 30 for Criteria 2 what should the aggregated Principle score for 'Good Feeding' be?" Alternatively, "if the Criteria score was 30 for Criteria 1 and 70 for Criteria 2 what should the aggregated score for 'Good feeding' be?" This was repeated for all of the Criteria aggregations that make up the remaining three Principles. Individual panel member scores were combined using simple arithmetic means for the elicited Principle scores. The exercise was also repeated in relation to the aggregation of Principle scores into a single welfare score. This enabled Choquet Integral weightings to be calculated which could then be used to weight the aggregation of Criteria scores into Principle scores and Principle scores into a single score for all of the policy case study scorings of the panel. The average scores elicited from the panel to calculate the Choquet weightings are shown in Annex 4. The same set of weightings were used for the six animal species considered. Separate weightings could have been elicited from the panel for each



species/animal type but those elicited in the WQ project showed that the weightings were not substantially between animal types.

The Choquet weightings were calculated by a process that involved selecting weightings that best explain the response scores provided by the experts. This selection is made based on the weightings that, when used within the Choquet Integral, minimize the sum of squared errors in the response scores given by the experts. The experts are assumed to use a Choquet Integral with these weightings but with some error. The weightings that result in the lowest sum of squared errors across all experts are then chosen as the final estimates of the weightings. This process is akin to using non-linear least squares, where the estimates are constrained to meet the conditions necessary to be valid weightings.

Mathematically, we have:

$x_{ij}$  the  $i$ th vector of scores of the given to the  $j$ th expert

$y_{ij}$  the response score of the  $i$ th expert conditional on  $x_{ij}$

$\phi$  the capacities that define the Choquet Integral

Let  $f(\phi, x_{ij})$  be the choquet integral of  $x_{ij}$  given capacities  $\phi$  then we model the process as equation 1:

Equation 1

$$y_{ij} = f(s, x_{ij}) + e_{ij}$$

The estimated capacities  $\phi$  are chosen so as to minimize (where the summation is with respect to all  $i$  and  $j$  - equation 2):

Equation 2

$$\sum_i \sum_j (y_{ij} - f(s, x_{ij}))^2$$

Python and Anaconda software were used to estimate the panel's Choquet weightings and to provide a computer program for aggregating scores from the expert panel's mean Criteria scores for each policy case study scenario. This program enabled rapid calculation of final Choquet-aggregated scores which could be undertaken in just a minute or two during the expert panel scoring meetings.

It is worth noting that WQ used Choquet Integral aggregate functions to aggregate Criteria scores into Principle scores but not Principle scores into a single score.

Resultant Choquet Integral weightings from the expert panel elicitation and the above calculation method, showed that in all cases the panel gave greater weight to low Criteria and Principle scores compared to high Criteria and Principle scores, although the degree to which this was the case varied depending on the specific Criteria and Principles considered. This demonstrates the general

belief that high scores in one Criteria or Principle cannot fully compensate for low scores in other Criteria or Principles.

### **Policy Case Studies**

Policy case studies were selected to include a range of topical possible (but hypothetical) policies to test the welfare assessment method with coverage across the six farm animal types and production systems. In particular, policies were chosen to include a range of different scenarios and welfare implications, including welfare impacts of different likely effects such as the numbers and types of animals affected, the duration and magnitude of impact for animals affected and the nature of the production system (e.g. relatively intensive or housed systems or more extensive systems).

The six policy case studies (15 scenarios) that the panel were asked to score were (i) phasing out of colony cages in egg production (ii) an increase in space for broiler chickens (from 39kg/m<sup>2</sup> to 30kg/m<sup>2</sup>) (iii) reduced use of farrowing crates or 'free farrowing' for indoor pigs (iv) reducing lameness in dairy cattle (to 5% within-herd prevalence) (v) reducing lameness in beef cattle (to 3% within-herd prevalence) and (vi) use of pain relief for lamb castration or no castration of lambs in sheep production. Each of the policy case studies also considered the status quo in the UK which considered current practice on farms across the nation. The latter provided base points from which to consider the policy changes outlined above. Thus, each of the policy case studies was comprised of a status quo scenario and either one or two policy change scenarios.

### **Results - welfare assessment**

The results of the expert panel's scoring of the policy case studies are shown in Annex 5. Note that Annex 5 presents arithmetic mean welfare scores alongside the Choquet weighted scores merely as comparison to show the important impact of the Choquet weightings on the welfare assessments. The arithmetic mean welfare scores of Criteria into Principles and then of Principles into a single score are not valid indicators of welfare and should not be used as such. Only the weighted Choquet values should be used for welfare assessment. Annex 6 contains the individual scores and comments of panel members.

A summary of the final (Choquet) single welfare scores for each policy scenario is shown in Table 2 (using the 0-100 scale where zero denotes extreme suffering for the animal and 100 denotes the highest achievable welfare possible).

Summaries of the expert panel's reasoning behind their scoring for each policy scenario, taken from the panel's (recorded) scoring discussions, are provided in Annex 6.

It can be seen from Table 2 that final welfare scores for the current (baseline) production case-study scenarios range from 27/100 (indoor pigs) to 64/100 (beef production). The more-intensive housed systems have lower scores than more extensive outdoor/grazing-based systems. Some policy scenarios have a relatively large increase in welfare score associated with them compared with others. For example, the difference of welfare score between colony cage egg production and free-range egg production is more than 20 points because of the major differences between the two systems which impact widely on welfare. Similar reasoning can be used for the difference between welfare scores for current use of farrowing crates for indoor pig production compared to free farrowing. Although indoor sows generally spend around 25% of their lives in crates, the crate is very

restrictive with negative welfare impacts across a number of welfare Criteria, especially those concerned with the ability to perform normal behaviours. In contrast, the use of pain relief for lamb castration in sheep production is only scored three points higher than current practice in the UK. In large part, this is because castration of lambs has most impact on the welfare of young lambs for a relatively short period of time and this represents a small part of sheep production and the whole life experiences of sheep and lambs from birth to slaughter.

**Table 2. Final Choquet Integral single scores for each policy scenario**

Species	Production System	Choquet Integral Score		Species	Production System	Choquet Integral Score
<b>Laying Hens</b>	Colony cage production	32.2		Indoor pig production	Current use of farrowing crates	26.8
	Barn production	43.8			Limited use of crates	34.3
	Free-range (not organic)	51.3			Free farrowing	46.6
<b>Sheep production</b>	Current with lamb castration	52.7		Dairy production	Current practice and lameness	43.1
	Castration with pain relief	55.6			Reduced lameness to 5%	56.1
	No castration	60.0				
<b>Beef production</b>	Current practice and lameness	58.5		Broiler production	Stocking density at 39kg/m2	37.9
	Reduced lameness to 3%	64.2			Stocking density at 30kg/m2	46.6

The reduction of lameness in dairy cattle (from the current average herd prevalence of around 30% to 5%) results in an increase in welfare score of 13 points in part because the reduced-lameness scenario would require a number of husbandry changes to achieve such a reduction in lameness and these would also benefit other areas of welfare impact for dairy cattle (such as comfort when resting). The reduction in lameness in beef cattle has a smaller increase in welfare score compared to dairy, in large part because the current (baseline) prevalence of lameness in beef animals is lower than that for dairy cattle. A reduction in stocking density for broiler chickens from a current legal maximum of 39kg/m2 to 30kg/m2 shows an increase in welfare score of nearly nine points (23%). The current baseline score takes into account the fact that a significant number of UK producers already stock below the legal maximum (e.g. Red Tractor at 38kg/m2). The score for the increase in space takes into account that the lower stocking density has most welfare impact in the last week or so of a circa six-week life for the broiler chicken when they are close to their maximum weight and size prior to slaughter, although by this time even 30kg/m2 allows only for restricted movement of birds.

## **Economic Valuation**

### **Approach and methods**

The economic valuation method uses a stated preference choice experiment survey approach developed from a previous study by Bennett et al (2012). The survey is directed at citizens who are the main food shoppers in their household in the UK. The survey is used to estimate households' willingness to pay based on the choices that survey respondents make from a series of choice questions. These choice questions show different combinations of animal welfare scores and increases in the weekly household food bill that respondents may be willing to pay for those scores. The focus on households is because most food shopping is done at the household level with food most often shared within a household. The main food shopper for the household is therefore most likely to be well-informed about the food consumption, purchases and preferences of the household generally.

A questionnaire was developed using the questionnaire from the Bennett et al (2012) study as a starting point. The development was undertaken over a series of months and involved initial exploratory research using four online small-scale surveys, six focus groups and eight verbal protocol interviews (where participants 'think aloud' while completing the questionnaire) to test various questionnaire designs and question formats (details provided below).

### **Exploratory research**

#### *Surveys*

Four online exploratory surveys were conducted using a commercial platform each with over 100 respondents selected at random from a national panel of people who have agreed to take part in online surveys administered by the commercial partner. The surveys were undertaken in parallel with the focus groups so that each was able to inform the other.

The first was a contingent valuation (CV) survey of 104 respondents designed mainly to explore the range of willingness to pay values that people had to improve the welfare scores of the farm animals involved in the study. A copy of the CV questionnaire is shown in Annex 7. In addition to questions very similar to those noted below for the main survey, the questionnaire contained four contingent valuation questions in place of the choice experiment questions used in later surveys. These asked for people's willingness to pay as an increase in their weekly food bill for their animal food products to come from animals with a 10-point and then a 20-point increase in welfare score for all six of the main farm species. This was followed by eight possible answers each showing a range of money amounts (e.g. £3-£5.99) from 0 to 'more than £20'. A similar two questions were repeated within the questionnaire but this time in the context of regulation agreed by government and the farming industry whereby all farm animals in the UK (in the six farm systems considered) would have to be kept at the increased 10 or 20 welfare score levels.

The findings showed that respondents to the survey would be willing to pay a mean of around £4.70 per week for a 10-point increase, using an approximation of people's willingness to pay as the middle value of the range that they had chosen. This value was only marginally more for a 20-point increase (£5.10) and marginally more again in the regulation scenario affecting the whole population of animals (£5.84 for the 20-point increase), with 87% of respondents having a willingness to pay of

less than £12 (96% of respondents gave a willingness to pay amount of under £12 for a 10-point increase).

The second, third and fourth surveys used the choice experiment method to elicit people's willingness to pay for higher welfare scores with sample sizes of 120, 105 and 107 respondents respectively. (The questionnaires can be found in Annex 8, 9 and 10 respectively). The questionnaires for these were very similar to that described below for the main survey (especially survey 4) but tested some (relatively minor) changes to the information statement presented to respondents and variants in some other questions. These surveys also contained a number of 'test' choice experiment questions to explore how people answered them and the rationality of their choices. Around 15-20% of responses to some test questions were not as expected. Such responses to the test choice questions together with changes in various aspects of the questionnaire design in the different surveys were also explored in the focus groups (see below) to better understand people's thinking when completing the questionnaire and help inform its further design.

Across the four surveys, 20-39% of respondents made some purchasing decisions because of their concerns about animal welfare, 58-85%, paid attention to the impact on their weekly food bill when making their choices to the willingness to pay questions, 67-86% had confidence in their answers, 73-89% understood the information presented to them and 68-89% liked the idea of farm regulation that affected all farm animals in the UK (highest in the CV survey for the confidence, information and regulation questions).

### *Focus groups*

Six focus group discussions were undertaken online with 7-8 people in each group (a total of 46 participants). Participants were recruited by a market research company from across the UK with a range of income levels and ages (and some other considerations for particular groups such as having one or two vegans or vegetarians). Main food shoppers for the household were targeted but some participants were not the main household food shoppers to gauge their potentially different perspectives. Two of the focus groups were female only and the four others were mixed sex. All female groups were chosen because female participants in particular can be inhibited in discussions involving some male participants. Each discussion lasted 1.5 hours and were led by a facilitator and supported by an assistant. Participants were paid a fee of £50 each for their participation.

The primary focus of the group discussions was the questionnaire design. Participants completed a specimen questionnaire (which varied from group to group in sequence with the exploratory surveys described above) within 24 hours prior to the discussions. This enabled the researchers to scrutinize participants' responses prior to each discussion. The discussions consisted of a 'warm up' phase discussing participants' attitudes to farm animal welfare and their purchasing behaviour regarding food animal products. This was then followed by going through the questionnaire question by question and asking whether they understood the question, whether they found it straightforward to decide their response etc. Particular attention was paid to people's understanding of the information statement and of the task involved in responding to the choice questions. The discussions were informal and free-flowing and the facilitator ensured that all participants contributed and that discussions were not dominated by a few individuals. All discussions were recorded and transcribed.

The main findings of the focus groups highlight that participants found the questionnaire interesting and understood the information presented to them (only one participant out of 46 said she did not really understand the questionnaire). A number of participants wanted to know more about the scoring system. Some wanted examples of what particular score levels would mean for the animals involved and how scores would be increased. Additional information was added to the information statement in the questionnaire in response to this. Most participants said they found the questions straightforward to answer including the choice questions, although some stated that the choice questions required additional thought to decide their preferred option. They found the layout of the choice questions clear and liked that the welfare levels that changed in each option were identified clearly by being in red which reduced the time and effort they had to devote to answering the questions.

From the discussions, it was clear that when completing the choice questions participants had a number of common approaches in relation to their thinking and decision making. For example, many participants paid most attention to the welfare scores of the products that their household most consumed and purchased. Some only paid attention to these scores and not to others. Most participants paid attention to the amount by which their weekly food bill would increase. A minority did not wish or could not afford to pay more for their food to increase animal welfare and had decided to choose option 1 (the 'no change' status quo option) for each of the choice questions presented to them.

Often participants stated that they had a weekly amount in mind that they could afford and would be willing to pay to improve the welfare of farm animals and so made their choice accordingly. This would sometimes mean that some of these people paid little attention to the welfare scores and chose their preferred option using the increase in their weekly household food bill as a guide to the likely welfare increase (i.e. assuming that the more they paid the better animals' welfare would be) or as a general proxy of quality. Indeed, when questioned why some participants had chosen a more expensive choice option for no increase in welfare scores (one of the 'test' choice questions), they responded that they had assumed that the option was superior in some other way because it was more expensive. These ways of thinking and approaches to answering the choice questions described by participants help to explain some unexpected choices that some people made that were observed in the exploratory surveys as well as the focus groups.

Transcripts of the focus group discussions are contained in Annex 11.

### *Verbal protocol interviews*

Verbal protocol interviews involve asking participants to share their thoughts in 'real time' by 'thinking out loud' and providing a running commentary of their thinking as they undertake a task. They can be very insightful as they provide spontaneous reactions and comments from participants rather than allowing post-rationalization. They are particularly useful to better understand people's thought processes and thinking when they are undertaking a task.

Eight verbal protocol interviews were undertaken with a range of different individuals. Verbal protocol interviews are where participants are asked to verbalize their thoughts (think out loud) as they are undertaking an exercise. In this context, individuals were asked to share their thoughts as they completed the questionnaire. The near-final version of the questionnaire was chosen to be

used for the interviews. Interviews lasted 30 minutes with a brief feedback discussion after they had completed the questionnaire and their commentary. Participants were mainly recruited by a market research company and were paid £25 each for their time (three were co-opted by the researchers and not paid).

The interviews identified which questions people spent most time on, which ones they found unclear or most cognitively demanding to answer and why they answered the questions the way they did. Generally, as in the focus groups, people found the questionnaire straightforward to answer although, again, a number noted that they needed to think carefully about their responses to the choice questions. Some identified the information statement as being rather long (which resulted in some editing to reduce the length where possible) although they found the information interesting and, in common with many focus group participants, wanted to know more about the welfare score and about the welfare of animals in the different production systems. After completing the questionnaire, some people stated that they would perhaps like to change one or two of their responses to the choice questions. One participant, who completed the questionnaire in a relatively very short period of time, said that he had only scanned the information statement and had not looked at the detail of the options in the choice questions.

Some example comments of participants made whilst they were completing the questionnaire show the kind of insights gained from the interviews: “Oh, there’s a lot of writing” (on seeing the information statement); “I like cows more than chickens. But it did say that there were more chickens. So, for the greater good do we save more chickens? Because £2 isn't a lot. I think I'm saving more chickens. Sorry, cows” (when deciding which option to choose from the three options in one of the choice sets).

Anonymised transcripts of the verbal protocol interviews are contained in Annex 12 together with a description of the task for participants.

### **Final survey design**

The final survey questionnaire design was decided following the testing and findings of the exploratory research. The final questionnaire consists of seven main sections which are:

1. Some socio-economic and demographic questions to allow quota sampling of the population such as whether they are the main food shopper for the household, their age, household income and place of residence. This helped to ensure a representative sample of the UK population.
2. Questions about what animal food products their household consumes and how much the household spends on food and different animal food products (i.e. beef, chicken, pigmeat, lamb, eggs, dairy products) per week. This not only provided information about household livestock product consumption patterns but also focused participant’s minds of this aspect of their food shopping and how much they spend on these items.
3. Some attitudinal questions about whether concerns about animal welfare influence their purchasing decisions and the extent to which they agree with certain statements (e.g. “I am concerned about the way farm animals are treated in the UK”). Those concerned about animal welfare are more likely to take the questions seriously and have a higher willingness to pay than those who don’t (other things being equal).

4. An information statement which tells respondents about a new method for assessing animal welfare and regulation agreed by government and industry to improve farm animal welfare (see below for the full statement). This was an important part of the questionnaire as it introduced and 'set the scene' for the choice questions which followed.

5. Twelve choice questions from which respondents must choose their preferred choice from three different options in each choice question (see below for a specimen choice question with example values). Responses to the choice questions were used to estimate people's willingness to pay for different welfare score levels for each animal type.

6. Three questions to check people's reasoning behind the choices they made. The first is an open text box that asks them to briefly state their reasoning. The second question asks whether they paid attention to particular aspects of the options presented to them in the choice questions when they decided which option to choose. The third question contains a series of statements to which respondents can agree or disagree to varying extents on a seven-point Likert scale (e.g. "I understood the information presented to me") to check how well they are likely to have completed the questionnaire.

See Annex 13 for a copy of the questionnaire showing all six of the different groups/blocks of 12 choice sets (only one block of choice sets is randomly allocated to each respondent).

Most important to the elicitation of people's willingness to pay are the information statement and choice questions. Below is a full version of the information statement.

---

### ***Animal welfare measurement and welfare score***

*Animal welfare scientists and veterinarians have developed a method for measuring the welfare of farm animals that takes account of the extent to which the needs and wants of the animal are met and results in an overall welfare score from 0 to 100.*

*The score accurately represents the welfare of the animal in terms of its freedom from hunger, thirst, discomfort, pain, injury, disease, fear and distress, and the extent to which the animal can express normal behaviours and has a happy and contented life.*

*A score of zero denotes extreme suffering whereas a score of 100 denotes the highest level of welfare that could possibly be achieved. The method covers the entire life of the animal from birth to slaughter and involves regular independent monitoring of farms.*

*Currently, in the UK (to the nearest 10 point mark):*

- laying hens have an average welfare score of 50/100*
- meat chickens 40/100*
- pigs 40/100*
- beef cattle 60/100*
- dairy cattle 40/100*
- sheep and lambs 50/100*



*Assume that the government and industry agree to farm regulation to improve the welfare of farm animals in the UK to a certain welfare score. **All** animals in the UK would have to be kept at this welfare score (or higher).*

*This higher welfare would incur additional production costs because animals would, for example, be better fed, have better housing and more space, better health and care, and more opportunities for normal social and other behaviours. These higher costs would result in more expensive meat, dairy, eggs and food with these ingredients, so that everyone will have to pay more for these food products, including you.*

*Your payment will be contributing to the higher welfare of farm animals throughout the UK.*

***Please choose out of the three options below your single most preferred option. If you do not like Choice 2 or Choice 3, choose Choice 1 which represents no change to current animal welfare levels and no change to your weekly food bill.***

*Remember that you have a limited budget and that more money spent on higher welfare food products means less money for you to spend on other things.*

*In the UK, around 3 million cattle & calves, 11 million pigs, 14 million sheep and lambs and 1 billion chickens are used for meat production, 2 million dairy cattle used to produce milk and 40 million laying hens used to produce eggs.*

---

Figure 2 is an example of a choice set with 12 similar choice sets being presented to each individual respondent in the questionnaire.

Of the three choice options shown in each choice question or 'set', the first option is always the status quo, where the welfare scores of the animal populations do not change and there is also no change to people's weekly household food bill.

The other two options both show changes to welfare score levels of only three of the six animal populations considered together with an increase to the household weekly food bill. The reason that only three welfare levels are changed in the second and third options of each choice set is because we found that the cognitive load of having more than three scores changing at one time was too high for many respondents. However, it is important that people make their choices in relation to all six of the major livestock systems in the UK and so all six species are shown.







The levels of farm animal welfare scores considered in the choice sets were 40, 50, 60, 70, 80 and 90 whilst the level of increases to people's weekly household food bills were £1, £2, £3, £4, £6, £8 and £10. These 'price' levels were chosen in the light of findings from the exploratory surveys and focus groups described above. An efficient choice set design (called a D-optimal design) was undertaken in relation to the statistical model to be used to estimate people's willingness to pay. This design contained different combinations of the above values with 72 different choice sets arranged into six groups of 12 choice sets. One out of the six choice set groups was then randomly allocated to each

respondent. There were three options in each choice set, one of which was always the status quo. None of the options showed a reduction in welfare score relative to the status quo level for each species/system (i.e. welfare score levels in options 2 and 3 either stayed the same as the status quo or increased). This is because people would not have a willingness to pay to reduce welfare scores below current status quo levels and such options within a choice set would not only be redundant (and inefficient) but would be likely to cause confusion for respondents.







**Figure 2. An example of a choice set**

Please choose **one** from the three choices shown below. Numbers represent welfare scores of the animals and products shown and the impact on your weekly food bill. **Welfare scores shown in red are the only ones that change in relation to current welfare scores in each choice.**







**Choice 1**

<input type="radio"/> <b>No change in weekly food bill or welfare scores</b>	 Eggs <b>50</b>	 Dairy (milk, butter, cheese etc.) <b>40</b>	 Pig meat (ham, bacon, pork etc.) <b>40</b>	 Lamb <b>50</b>	 Chicken meat <b>40</b>	 Beef <b>60</b>
--	--	---	--	---	--	--

**Choice 2**

<input type="radio"/> <b>£2 increase in weekly food bill (£104 per year)</b>	 Eggs <b>80</b>	 Dairy (milk, butter, cheese etc.) <b>60</b>	 Pig meat (ham, bacon, pork etc.) <b>40</b>	 Lamb <b>90</b>	 Chicken meat <b>40</b>	 Beef <b>60</b>
--	--	---	--	---	--	--

**Choice 3**

<input type="radio"/> <b>£10 increase in weekly food bill (£520 per year)</b>	 Eggs <b>90</b>	 Dairy (milk, butter, cheese etc.) <b>60</b>	 Pig meat (ham, bacon, pork etc.) <b>40</b>	 Lamb <b>50</b>	 Chicken meat <b>40</b>	 Beef <b>90</b>
---	--	---	--	---	--	--

The questionnaire was administered online using a commercial survey platform. The platform uses panels of citizens prepared to take part in online surveys. Quota sampling ensured a stratified random selection of participants to be representative of the UK population across the four countries of the UK (England, Northern Ireland, Scotland and Wales). The sample was stratified according to age, household income, region of habitation in the UK, ethnicity and education.

The survey was piloted (on 50 respondents) prior to full launch and the collected data scrutinized as a final test to check that the questionnaire was working as intended. The survey was then fully launched on September 3<sup>rd</sup> 2024 and completed on October 15<sup>th</sup> 2024 with a total of 3,013 completed questionnaires obtained. Annex 14 shows the quota levels in the survey.

### Method of analysis of willingness to pay

To analyse the discrete choice experiment (DCE) data, a Hierarchical Bayesian Logit (HBL) was employed (Balcombe et al. 2016). A HBL is a flexible approach that allows for a continuous distribution of preferences across the population. When implemented in ‘willingness-to-pay space’ the model also affords the introduction of prior information regarding reasonable bounds upon the distributions of respondents’ willingness to pay (wtp). Popular alternatives to this model include classical approaches such as the Mixed Logit and the Latent Class models. However, the HBL can deliver similar results to the classical Mixed Logit but offers the opportunity to use relatively ‘weak’ information such as extreme bounds for the distribution of normally distributed parameters that can greatly improve estimates under some circumstances. Our choice of model also reflects our belief that respondent heterogeneity is best modelled as a continuous distribution rather than treating respondents as being drawn from a set of groups which is implied if employing a Finite Mixture/Latent Class model.

Our model specification is formally defined as follows. Let  $x_{ijs}$  denote a  $k \times 1$  vector of attributes (a payment and set of welfare scores) from the DCE presented to individual  $j$  ( $j=1, \dots, J$ ) in the  $i$ th option ( $i=1, \dots, I$ ) of the  $s$ th choice set ( $s=1, \dots, S$ ). Next, let  $U_{ijs}$  be the utility that individual  $j$  attains from  $x_{ijs}$ . Given these definitions, it then follows that an individual  $j$  is assumed to receive linear utility from the  $i$ th choice in the  $s$ th choice set. Consequently, the utility function (Equation 3) is of the form:

Equation 3

$$U_{ijs} = V_j(X_{ijs}) + e_{ijs}$$

where  $V_j(x_{ijs})$  is the systematic utility that individual  $j$  obtains from the vector  $x_{ijs}$ . The error term  $e_{ijs}$  is assumed to be extreme value (Gumbel) distributed, independent of  $x_{ijs}$  and uncorrelated across individuals or choices. It then follows that the probability of choosing option  $i$  for the  $j$ th person from the  $s$ th choice set (Equation 4) is:

Equation 4

$$p_{ijs} = \frac{e^{V_j(X_{ijs})}}{\sum_i e^{V_j(X_{ijs})}}$$

As is becoming common in the DCE literature, we estimate our models in what has been termed wtp space. The reason for adopting this approach is that it can significantly reduce the instability associated with wtp estimates recovered from preference space (Balcombe et al., 2010). It also means that model parameters are directly interpretable as wtps. From a Bayesian perspective, DCE models usually require some level of informativeness in the priors. Having the parameters representing wtps means that formulating sensible priors is far easier in wtp space since very often we will have some prior idea of the likely values of these parameters. However, in large datasets (such as the one here) these priors will have little effect since they will be dominated by the information in the data.

The systematic utility component we employ in this paper is as follows in Equation 5:

Equation 5

$$V_j(x_{ijs}) = \exp(\beta_{0,j}) \left[ -\text{price}_{ijs} + \beta_{1j}\text{henslaying}_{ijs} + \beta_{2j}\text{dairycattle}_{ijs} + \beta_{3j}\text{pigs}_{ijs} \right. \\ \left. + \beta_{4j}\text{sheep}_{ijs} + \beta_{5j}\text{broilers}_{ijs} + \beta_{6j}\text{beefcattle}_{ijs} + \beta_{7j}\text{SQ}_{ijs} \right]$$

where  $\beta_{1j}, \dots, \beta_{6j}$  represent wtp parameters for the  $j$ th individual for the welfare (measured as a score from 0-100), whereas  $\beta_{7j}$  is the wtp to avoid the Status Quo (the existing levels of welfare).

The prior mean wtp was set equal to zero with standard deviation of £10 for a 10 point per annum improvement in welfare for each species, using a positive half normal distribution (such that wtps were non-negative). The individual specific wtps were then specified as hierarchically normally around the estimated wtp means with standard deviations that followed a Gamma distribution.

The specification above is a standard approach used in choice experiments, which almost universally assume that the wtp is constant over the range of prices. However, since we believe that in this instance people may have diminishing wtp over the score range we specified a function that allowed for decreasing wtp as the increments above the status quo level become higher. Formally, we specified the function in Equation 6:

Equation 6

$$V_j(x_{ijs}) = \alpha \left( -p_{ijs} + \sum_k \beta_{k,j} \left( x_{ijs,k} - \frac{1}{2} \rho \left( \frac{(x_{ijs,k} - x_k^{sq})}{(90 - x_k^{sq})} \right)^2 \right) \right)$$

where  $x_{ijsk}$  denotes the  $k$ th element of  $x_{ijs}$  and  $x_k^{sq}$  is the status quo level of the  $k$ th attribute. This model introduces the parameter  $\rho$  which at 0 gives the standard model and at 1 implies that people's wtp for a score of 90 (the highest level of welfare we offer in the experiment) is zero. Given an estimate of  $\rho$ , the willingness to pay for the  $k$ th attribute at the level is seen in Equation 7:

Equation 7

$$\text{WTP}_{k,j} = \beta_{k,j} \left( 1 - \rho \left( \frac{(x_k - x_k^{sq})}{(90 - x_k^{sq})} \right) \right)$$

This method allows us to estimate marginal wtps across the range of animal welfare scores (above the status quo values).

## Results

### *Characteristics of the sample – descriptive statistics*

It should be noted that a number of characteristics of the sample were largely determined by the quota requirements.

Thus, in relation to household income (before taxes and including income from investments, pensions, benefits etc), decile income bands from the Office of National Statistics (ONS) data, were used for quota sampling and so around 10% of the sample were contained in each of the 10 income bands which were: 0-£19,999, £20,000-£27,999, £28,000-£34,999, £35,000-£41,999, £42,000-£49,999, £50,000-£58,999, £59,000-£68,999 (only 8% of the total sample due to relaxing quota requirements as explained above), £69,000-£83,999, £84,000-£144,999 and £145,000 and above.

Table 3 below shows the percentage of respondents according to age category.

**Table 3 Age of respondents (% respondents in each age category)**

Age category (yrs)	% of respondents in the sample
18 - 24	9
25 - 34	17
35 - 44	17
45 - 54	15
55 - 64	17
65 - 74	14
75 and over	10

*Numbers may not add to 100% due to rounding*

Region of habitation was subject to quota requirements, to ensure a representative sample, with Scotland, Wales and Northern Ireland having higher quotas than in the general population to ensure sufficient sample sizes to undertake some analyses by country within the UK. Table 4 shows the percentage sampled in each region.

**Table 4 UK region of habitation of respondents (% of respondents)**

Region of the UK	% of respondents in the sample
North-East England	2
North-West England	9
Yorkshire & the Humber	7
East Midlands of England	6
West Midlands of England	7
East of England	8
London	12
South-East England	12
South-West England	7
Wales	10
Scotland	10
Northern Ireland	8

Table 5 below shows the percentage of respondents according to their highest educational qualification.

**Table 5 Highest educational qualification of respondents (% of respondents)**

Highest educational qualification	% of respondents
No formal qualifications	3
GCSE or equivalent	20
Apprenticeship	3
A level or equivalent	23
HNC, HND, Bachelor's degree, post-graduate degree	49
Other qualification	2

In terms of the gender of respondents, 65% were female and 35% male. Seventy three percent of respondents stated that they did all of the food shopping for the household and 27% did most of the food shopping (the survey sampled only those that did most or all of the food shopping for the household but with no quota requirement as to whether they did all or most of the food shopping).

In terms of the ethnicity of respondents to the survey, 86% were White, 8% Asian or Asian British, 3% Black or Black British, 2% Mixed or Multiple Ethnic Group and just 1% of some other ethnic group.

Seventy-two percent of households represented in the survey were one family households, 18% were one person households, 7% multi-family households and 3% households of unrelated adults.

Mean household size was 2.2 adults (range 0-11 people) and 0.7 children under 16 (range 0-11 children).

Sixty-one percent of respondents were employed, 22% retired, 7% self-employed, 6% not working, 3% unemployed (but available for work) and 3% were students.

The median time for completion of the questionnaire by respondents was 7.3 minutes.

### *Food consumption and attitudes of the sample*

Table 6 below shows household expenditure on food each week.

**Table 6 Household expenditure on food (£/week)**

<b>Household expenditure on food (£/week)</b>	<b>% of households</b>
<£50	13
£50 - £99	37
£100 - £149	24
£150 - £199	10
£200 - £249	6
£250 - £299	4
£300 and over	5

Ninety-four percent of respondents eat eggs, 94% eat chicken, 93% eat (at least some) dairy products, 88% eat beef, 74% eat pigmeat, 69% eat lamb, 64% eat other animal products (e.g. fish) and 1.5% do not eat animal products.

Sixty-six percent of respondents stated that concerns about animal welfare influenced their purchasing decisions.

Table 7 below shows household expenditure on specific animal products.

**Table 7 Household expenditure on the main animal food products (£s/week)**

<b>Weekly household expenditure</b>	<b>% of households</b>					
	<b>Chicken</b>	<b>Beef</b>	<b>Lamb</b>	<b>Pigmeat</b>	<b>Dairy</b>	<b>Eggs</b>
£0	4	12	39	20	1	5
£0.01 - £4.99	17	25	21	32	24	64
£5 - £9.99	29	26	16	23	30	16
£10 - £14.99	19	16	9	10	20	7
£15 - 19.99	11	6	4	5	10	5
£20 - £29.99	9	5	4	5	7	2
£30 - £39.99*	5	4	4	3	4	1*
£40 - £50	3	3	2	1	2	
More than £50	3	3	2	1	1	

\* More than £30 per week in the case of eggs. Numbers may not add to 100% due to rounding

Most respondents agreed that food products from animals with high levels of welfare are healthier (68%), taste better (64%) or are better for the environment (73%). This shows that people generally perceive foods coming from animals with higher welfare as being of higher quality generally. Fifty-six percent agreed that they feel well informed about how farm animals are treated and 62% are concerned about the way farm animals in the UK are treated.

After completing the willingness to pay choice questions, respondents were asked to briefly explain the reasoning behind their choices. This provided evidence of people having thought about the choice questions and that they have reasons behind their responses. Respondents gave a range of

responses but the rationale for most respondents is encapsulated by the quotes below from three respondents:

“I was willing to go up to a £4 per week increase, but no further. Then I was looking at the meat type I personally eat most often.”

“An increase above £5 starts to seem a lot compared to what I pay now, seeing as prices are already high these days. I wish the increase in amount I am willing to pay would increase the quality of life for all the farm animals, but if it can't I'm tending to favour the option that increases the quality of life for the larger animals, thinking that having bigger brains they might notice the increase in quality of life more? or, selfishly, the ones I eat the most eg chicken.”

“Purely based on price. I'm limited by financial constraints - I chose the option I could afford, which is always going to be less than the outcome I desire.”

Respondents were asked which features they considered when making their choices in the choice questions (from a list presented to them). Seventy-one percent said they considered the increase in their food bill, 60% considered chicken welfare levels, 54% dairy cattle welfare levels, 52% beef cattle welfare levels, 49% hen welfare levels, 43% pig welfare levels and 35% considered sheep welfare levels.

Eighty-three percent of respondents stated that they had confidence in their responses to the choice questions. Eighty-nine percent stated that they understood the information presented to them (e.g. in the information statement). Eighty-five percent of respondents agreed that we have a moral duty to safeguard the welfare of animals, 86% liked the idea of regulation to improve the welfare levels of all farm animals while 80% would like to see the welfare scoring system on labels in food stores. Seventy-four percent of respondents agreed that they trusted the animal welfare scoring method that had been described to them and 74% also trusted that farms would be properly monitored.

### People's willingness to pay to increase animal welfare scores

Model results using the 'standard' approach for analysing choice experiment data are shown in Table 8a.

**Table 8a. Estimates of mean annual willingness to pay for a 10-point increase in welfare score for each species/system (£s/year)**

	Laying hens	Dairy cattle	Indoor pigs	Sheep	Broilers	Beef cattle
<b>Mean</b>	31.41	36.64	29.74	28.24	32.75	31.65
<b>Std</b>	0.94	0.80	0.74	0.86	0.79	1.10
<b>Min</b>	27.45	33.66	27.10	25.17	29.67	27.77
<b>25%</b>	30.74	36.10	29.24	27.67	32.21	30.90
<b>50%</b>	31.39	36.65	29.73	28.24	32.73	31.64
<b>75%</b>	32.04	37.17	30.21	28.82	33.30	32.38
<b>Max</b>	35.20	39.48	33.20	31.42	35.44	35.85



Distribution of wtp estimation across the iterations (30,000) shown by the 25%, 50% and 75% values (in addition to the 'std' which is the standard distribution) can, in Bayesian models, be interpreted in a similar way as confidence intervals in classical models. The values are relatively tightly distributed suggesting a good level of confidence in the model results. 'Min' shows the minimum value for mean wtp from the iterations and 'max' the maximum value and these can be considered analogous to a 99% confidence interval. The model also has a McFadden (Pseudo) R Squared of 0.233. The McFadden R Squared is a common measure of 'goodness of fit' of Bayesian models and is analogous to the R Squared measure in a regression model. A McFadden R squared above 0.2 is considered to signify a model with good fit.

The results above show a range of mean wtp across the species/systems. Values for smaller or greater changes in welfare score than the 10-point increase values shown in Table 8 can be calculated pro-rata (e.g. wtp for a one-point increase is merely one tenth of values shown in Table 8). The range of mean wtp values, when comparing across the species, varies from a low of around £28/yr for sheep to nearly £37/yr for dairy cattle. When thinking about differences in these values a number of considerations should be factored in. These include the different status quo 'starting point' scores of the different species presented to respondents which range from 40 for broilers, pigs and dairy cattle to 50 for laying hens and sheep/lambs to 60 for beef cattle. In addition, the different species and their products have different levels of importance to respondents across the sample often depending on their household consumption of the products in question. For example, lamb was less commonly consumed by households than other products, whilst dairy, chicken and eggs were most commonly consumed. Respondents may also have had a preference for certain species, such as cattle, seeing their welfare as of more importance than some others (e.g. a quote from one of the verbal protocol interviews was "...I like cows more than chickens ..." when making a decision about which option to choose for one of the choice sets in the survey questionnaire) or may be concerned about their perception of relatively low status quo welfare scores for some species (".. that's awful that dairy and eggs are so low" quote from the same verbal protocol interviewee when looking at the status quo welfare scores).

Mean wtp levels across the species show that people would be willing to pay the equivalent of around £190 per household per year for a 10-point welfare score increase for all six species. Mean wtp for a 10-point increase in welfare of the relevant species as a percentage of respondents' household expenditure on the products of that species is 4.5% in the case of chicken meat, 6.2% for beef, 6.3% for dairy, 6.7% for lamb, 7.1% for pig meat and 11.2% for eggs.

Table 8b shows results from the estimation of marginal willingness to pay from the advanced/extended Bayesian model.

Mean is the mean wtp (£s/household/year), from all iterations of the model, for a one-point increment in welfare score from the status quo value of the animal type concerned. For example, the table shows a wtp of £5.18/household/year to increase the welfare of dairy cattle in the UK from the status quo value of 40 to 41.

**Table 8b. Extended model wtp (£s/household/yr) results**

	Laying hens	Dairy cattle	Indoor pigs	Sheep	Broilers	Beef cattle
<b>Mean</b>	4.44	5.18	4.27	4.04	4.56	4.40
<b>Std</b>	0.19	0.17	0.17	0.16	0.16	0.21
<b>Min</b>	3.74	4.61	3.70	3.44	3.99	3.75
<b>25%</b>	4.31	5.06	4.16	3.93	4.45	4.25
<b>50%</b>	4.43	5.18	4.27	4.04	4.56	4.40
<b>75%</b>	4.57	5.29	4.38	4.15	4.66	4.53
<b>Max</b>	5.12	5.75	4.84	4.60	5.17	5.09

Table 9 below shows additional results from the estimation of marginal willingness to pay from the extended Bayesian model for successive one-point increases in the welfare score, as snapshots at different points further from the status quo levels for each animal type.

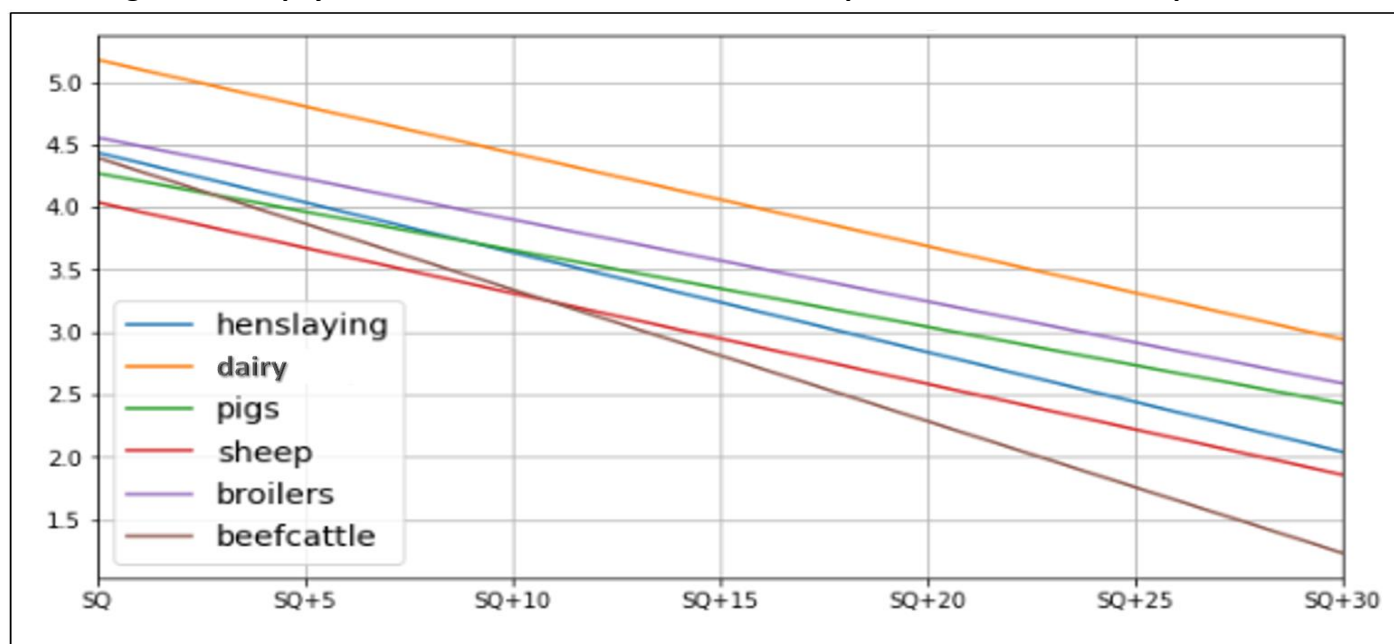
**Table 9. Marginal wtp (£s) for a one-point increase in welfare score at various points from the status quo (SQ)**

	SQ	SQ+1	SQ+5.00	SQ+10.00	SQ+15.00	SQ+20.00	SQ+30.00
<b>Hens laying</b>	50	4.44	4.04	3.64	3.24	2.84	2.04
<b>Dairy</b>	40	5.18	4.81	4.43	4.06	3.69	2.94
<b>Pigs</b>	40	4.27	3.96	3.66	3.35	3.04	2.43
<b>Sheep</b>	50	4.04	3.68	3.31	2.95	2.59	1.86
<b>Broilers</b>	40	4.56	4.23	3.90	3.58	3.25	2.59
<b>Beef cattle</b>	60	4.40	3.87	3.34	2.82	2.29	1.23

For example, in the case of laying hens, mean marginal wtp is estimated as £4.44 per year for a one-point welfare score increase for laying hens from a status quo score of 50 to a score of 51. This is somewhat higher than the estimate of mean wtp from the fixed wtp model across the full score range (presented previously in Table 8a), showing that people have a higher marginal wtp at lower levels of welfare score. This is borne out by the remaining marginal wtp estimates with, for example for laying hens, a value of £4.04 per year for a one-point increase from welfare score 55 (SQ+5) to 56, £3.64 for a one-point increase from 60 to 61, £3.24 for 65 to 66, £2.84 for 70 to 71 falling to £2.04 for 80 to 81.

A similar reduction in marginal wtp is seen for the other animal types, showing diminishing marginal wtp as the welfare score increases. This can be clearly seen graphically in Figure 10 below which shows the same results as those summarized in Table 9.

**Figure 10. Wtp/yr for a 1% increase in welfare score for points above the status quo**



The rate of decline of marginal wtp as welfare score increases is determined by the value of  $\rho$  (described in the 'method of analysis of willingness to pay' section above) which is estimated from the survey data. The value of  $\rho$  is between 0 and 1 (where 0 would mean that there is no reduction in marginal wtp as the welfare score increases and 1 means that marginal wtp declines to zero).  $\rho$  is estimated from the data at 0.72.

The advanced wtp estimation model has a McFadden R squared of 0.256 which is an improved 'goodness of fit' compared to the 0.233 of the basic model. Moreover, the advanced model estimates marginal wtp which is a more precise measure of wtp as welfare score increases compared to a mean wtp value applied to the whole score range (i.e. from status quo welfare score levels to a score of 90), and is more compatible with the expectation of economic theory of diminishing marginal consumer benefit.

Therefore, the advanced model estimation results are most appropriate to be used for policy analysis. In order to use these results, tables have been produced which show both the marginal and cumulative wtp benefit values of increases in the welfare scores of each of the animal types. Annex 15 contains the 'wtp valuation tables' which show marginal wtp values for each one-point increment in welfare score, together with cumulative values of the one-point increments.

### **Use of results for economic valuation of the benefits of policy scenarios**

This section considers each of the six-policy case studies and uses the welfare scores and score valuation tables to provide estimates of the economic value of policies to improve the welfare of animals. Summary benefit calculations for each of the policy scenarios considered are shown in Table 10.

**Table 10. Benefit calculations for each policy scenario<sup>(a)</sup>**

Policy change scenario	Welfare score change	WTP (£s per hh per yr)	UK total benefit (£M)	WTP (£) per unit of meat, milk or eggs	Retail price range (b) (£ per unit)	Unit
Broiler stocking density 38kg/m <sup>2</sup> (c) to 30kg/m <sup>2</sup>	39 to 47	35	997	0.92	3 to 18	per bird
Dairy cattle lameness reduced to 5%	43 to 56	59	1664	0.11	0.64 to 2	per litre milk
Beef cattle lameness reduced to 3%	58.5 to 64	24	669	0.74 (d)	5 to 60 (e)	per kg beef
Sheep: lamb castration with pain relief	53 to 56	11	320	1.12 (d)	8 to 30	per kg lamb
Sheep: no castration of lambs	53 to 60	25	717	2.51 (d)	8 to 30	per kg lamb
Indoor pigs: farrowing crates limited to 1 week	27 to 34	30	492 (f)	0.9 (d)	5 to 15	per kg pork
Indoor pigs: no use of farrowing crates	27 to 47	84	1386 (f)	2.52 (d)	5 to 15	per kg pork
Laying hens: colony cage to barn	32 to 44	53	315 (h)	0.13	0.13 to 0.66	per egg
Laying hens: colony cage to free range (g)	32 to 51	84	496 (h)	0.2	0.13 to 0.66	per egg

(a) Shows (i) welfare score due to policy change scenario (rounded to nearest whole number) (ii) household WTP for the score change (iii) total UK benefit (household WTP x 28.4M UK households) (iv) WTP per unit of animal output and (v) current retail prices per unit of animal food products for context.

(b) Retail prices were accessed between 12/11/2024 and 19/11/2024 from UK retailer websites.

(c) Most UK broilers are kept at Red Tractor assurance scheme 38kg/m<sup>2</sup>. (d) Carcase weight

(e) In-store packs ranging from minced beef to fillet steak. (f) Assumes that 58% of UK sows are kept in farrowing crates.

(g) Not organic. (h) Based on 21% of hens kept in colony cages in the UK.

A full description of the calculations and reasoning involved in each of the policy scenario benefit calculations are presented further below and the wtp valuation tables in Annex 15.

Note that some policy scenarios may involve starting-point welfare scores below the status quo values presented to respondents in the survey. This is because the status quo values used represented current welfare levels in the UK population generally (rounded to the nearest 10 points to make comparison of scores cognitively easier for survey respondents). However, some policy scenarios considered specific production systems only, such as colony cage egg production. Colony cage egg production has a lower welfare score than free-range and barn egg production which represent the majority of hens kept and eggs produced. Thus overall, across all hens in the UK, the welfare score is much higher than for those used in colony cage production.

It might be expected, from economic theory, that marginal wtp values to improve welfare scores might increase as starting point welfare levels decline below the status quo levels (these can be calculated by extrapolation using the marginal wtp rates shown in the wtp valuation tables). However, because the survey did not elicit wtp values from people for welfare scores below the status quo levels in the survey, the benefit calculations presented below do not use extrapolated values but instead take a parsimonious approach and use the lower status quo +1 values for all welfare score levels below the status quo values.

### Broiler case study

This involves two broiler maximum stocking density scenarios, (i) the current legal maximum of 39kg/m<sup>2</sup> with welfare assessment score of 37.9 and (ii) a maximum stocking density of 30kg/m<sup>2</sup> with welfare score 46.6.

Benefit valuation of an increase of score from 38 to 47 comprises two components. (i) From 38 to 40 the valuation tables use a value of £4.56/household/yr for a one-point increase in welfare score which for a two-point increase is  $£4.56 \times 2 = £9.12/\text{household/yr}$ . (ii) From 40 to 47 the valuation tables give a cumulative value of £30.54/household/yr which added to the £9.12 is £39.66 in total. This is multiplied by the number of UK households of 28.4 million (ONS (2024) data for 2023) = a benefit of £1,126M/yr based on people's willingness to pay assuming all 1,100 million broiler chickens produced per year in the UK are currently stocked at 39kg/m<sup>2</sup> and move to being stocked at 30kg/m<sup>2</sup>. This is the equivalent of a little over £1 per bird extra for the increase in welfare.

Interpolation from the welfare scores for stocking densities between 39kg/m<sup>2</sup> and 30kg/m<sup>2</sup> can be used to calculate welfare scores for any stocking density between the two values.

For example, the Red Tractor standard of 38kg/m<sup>2</sup> is met for over 95% of meat produced in the UK. A stocking density of 38kg/m<sup>2</sup> would have a welfare score of 39 (rounded and using interpolation between welfare scores for 39kg/m<sup>2</sup> and 30kg/m<sup>2</sup>). Using the valuation tables, an increase in welfare score from 39 to 47 has a wtp of (i) £4.56/household/yr to move from a score of 39 to 40 and (ii) £30.54/household/yr to move from a score of 40 to 47. This gives a total wtp of  $£4.56 + £30.54 = £35.10/\text{household/yr}$  which multiplied by 28.4 UK households = £997M which divided by 1.1 billion broiler chickens is just over £0.90 per bird. (This is an approximate value as some Red Tractor producers will have stocking densities below the Red Tractor standard whilst some of the 5% of producers not members of Red Tractor may have stocking densities in excess of the standard.)

The above analysis can be considered in the context of current UK retail milk prices paid by consumers. Current prices for fresh whole indoor-raised chicken are £3-£5 per bird (e.g. Tesco, Sainsbury) while free-range chicken varies from £6-£13 (e.g. Sainsbury, Asda) and more for organic (e.g. £18 per bird at M&S). The above wtp amounts calculated would appear to be well within the range of current prices paid by consumers for chicken with higher welfare attributes.

### **Pig case study**

(i) current use of farrowing crates (applies to indoor pigs only) with score 26.8 (ii) short-term use of farrowing crates with score 34.3 and (iii) free farrowing with score 46.6.

We use the current welfare score of 26.8 rounded to 27 as the baseline.

This gives two scenarios to value (i) an increase in welfare score from 27 to 34 and (ii) an increase in welfare score from 27 to 47 (rounded to nearest whole numbers).

The status quo score for pigs that was used in the valuation survey was 40. Using the welfare score valuation tables, we therefore need to apply a marginal willingness to pay value of £4.27 for each one-point increase in welfare score from 27 to 40.

(i) For an increase in welfare score from 27 to 34 (a move from current use of farrowing crates to only short-term use of crates), this gives a wtp of  $7 \times 4.27 = £29.89/\text{household/yr}$  which multiplied by 28.4M UK households = £849M/yr assuming all pigs move from an average welfare score of 27 to score 34.

(ii) For an increase in welfare score from 27 to 47 (a move from current use of farrowing crates to free farrowing), we apply the wtp of £4.27 for each one-point increment up to 40. This gives a wtp of  $13 \times 4.27 = £55.51/\text{household/yr}$ . We then need to calculate the cumulative wtp to move from 40 to

47 from the valuation table which is £28.60/household/yr. Adding these two wtp calculations together gives the wtp for moving all UK pigs from welfare score 27 to welfare score 47 which is  $£55.51 + £28.60 = £84.11$  which multiplied by 28.4M UK households = £2,389M/yr assuming all pigs move from an average welfare score of 27 to score 47.

However, although around 11 million pigs are farmed in the UK each year, producing around 948M kg of meat (carcase weight, AHDB, 2024) not all breeding sows are kept in farrowing crates because around 42% of breeding herds/sows are kept outdoors (out of a total of around 340 thousand female pigs in the UK breeding herd; AHDB, 2024), although the vast majority of pigs are reared indoors and nearly all pigs are finished indoors. It is therefore difficult to estimate the total benefit value of changes to the use of farrowing crates since fattening/finishing pigs are involved as well as sows. A crude (and likely conservative) calculation would be to assume that only 58% of all pigs are affected by changes in the use of farrowing crates such that the final benefit calculation using the figures derived above would be (i)  $£849M \times 0.58 = £492M$  for short-term use of farrowing crates for indoor pigs and (ii)  $£2,389M \times 0.58 = £1,386M$  for free farrowing.

Again, if we assume that 58% of pig meat is affected, the benefit calculations above equate to a little under £0.90/kg on the price of pig meat (carcase weight) for the increase in welfare from score 27 to 34 for the short-term use of farrowing crates and £2.52/kg on the price of pig meat (carcase weight) to move from the current welfare score of 27 to that of 47 associated with free farrowing.

To put the above analysis into context, retail (supermarket) pig meat prices currently paid by consumers vary from round £5/kg (e.g. Tesco pork shoulder, Sainsbury loin joint, Aldi leg joint) to £9/kg (e.g. Aldi British pork loin, Waitrose pork belly, Sainsbury pork loin steaks) to £15/kg and more for free-range/organic pig meat or for processed products such as gammon (e.g. Waitrose gammon, Waitrose organic pork joint) and substantially higher prices for ham (e.g. £20/kg for Waitrose ham, £27/kg for Tesco finest ham).

Thus, estimated wtp for higher welfare pig meat would appear to be well within the current prices paid for UK-produced pig meat marketed as higher welfare (such as free range), even allowing for the lower weight of fully butchered products in the store compared to carcase weight.

### **Dairy case study**

The expert panel scored two dairy scenarios; (i) current dairy production with 30% within-herd prevalence of lameness (mobility scores 2/3) with score of 43.1 and (ii) lameness reduced to 5% prevalence with score 56.1. The analysis below assumes that the prevalence of lameness is reduced to 5% across all dairy herds on average in the UK.

The scenarios outlined above require a welfare benefit valuation for an increase in welfare score from 43 to 56 (rounded to the nearest whole number).

Using the welfare score valuation table showing cumulative wtp values, the wtp for increasing the welfare of animals in dairy production from 43 to 56 is £58.61/household/yr (i.e. £73.93 - £15.32 on the cumulative valuation table). A wtp of £58.61/household/yr multiplied by 28.4M UK households = £1,664M.

In the context of dairy production, there are around 2.65M dairy cows in the UK December 2023; AHDB, 2024) which is £628 per cow/yr wtp to increase dairy welfare from a welfare score of 43 to 56.

Annual UK milk production (milk sold to dairies) was 14,866 million litres in 2023 (Defra, 2024a). This gives a wtp of 11.2p/litre (£1,664M divided by 14,866M). Around 40% of milk availability is liquid milk with the remainder used to produce higher value products such as cheese, butter etc.

To put the above analysis into context, the retail price of milk varies from £0.64/litre (generally for large two or four litre packs) to over £2/litre (e.g. Waitrose organic milk in a smaller container). Wtp to increase the welfare of dairy cows of 11.2p/litre is the equivalent of a 17.5% increase to the cheapest (£0.64/litre) milk (potentially increasing the price to around £0.75/litre) or a 6.8% increase to premium milk such as Waitrose Duchy organic (£1.65/litre). The wtp per litre of milk therefore appears to be well within the price range found for premium milk compared to more basic milk products. The potential percentage increases in the price of high value dairy products, such as cheese, would be somewhat lower than for liquid milk.

### **Sheep case study**

This case study comprises three scenarios that were scored by the expert welfare assessment panel: (i) current production with lamb castration and score 52.7 (ii) castration with pain relief with score 55.6 and (iii) no castration with score 60.0

The current welfare score (53 rounded) is used as a baseline and the benefit in terms of wtp is calculated for (i) an increase from welfare score 53 to 56 (rounded) for castration with pain relief and (ii) an increase from score 53 to 60 for no castration.

From the cumulative wtp value table, the wtp for score increase from 53 to 56 is £23.15 - £11.90 which = £11.25 per household/yr. The wtp for score increase from 53 to 60 is £37.13 - £11.90 from the cumulative wtp score table which = £25.23 per household per year.

Multiplied by the number of UK households of 28.4M = (i)  $28.4M \times £11.25 = £319.5M/yr$  to increase welfare score from 53 to 56 and (ii)  $28.4M \times £25.23 = £716.5M/yr$  to increase welfare score from 53 to 60.

UK lamb production was 286,000 tonnes in 2023 (carcase weight; AHDB, 2024). The 'castration with pain relief' scenario equates to a wtp of £1.12/kg of sheep meat and the 'no castration' scenario equates to a wtp of £2.51/kg of sheep meat.

Current prices of lamb in supermarkets vary greatly depending on cut, from around £8/kg (e.g. Iceland frozen chump chops) to leg of lamb at around £15/kg (e.g. Tesco) up to around £30/kg (e.g. Waitrose lamb leg steaks at £28/kg and Riverford organic chops at £29/kg). On the mid-priced lamb of around £15/kg, the 'castration with pain relief' valuation equates to a potential increase of 7.5% to the price and the 'no castration' calculation equates to a 16.7% increase to the price.

### **Beef cattle case study**

The expert panel scored two beef cattle scenarios (i) current production and lameness levels with score 58.5 and (ii) lameness reduced to 3% prevalence with score 64.2

This requires valuation of an increase in beef cattle welfare score from 58.5 to 64 (rounded). Using the wtp score value table, increasing score from 58.5 to 60 has a wtp of  $1.5 \times £4.40 = £6.60$  and then from 60 to 64 = £16.97 which added together = £23.57/household/yr.

£23.57 multiplied by 28.4M UK households gives a wtp for UK households of £669M/yr.

Putting this into context, in 2023, just over 2 million prime cattle were slaughtered for beef from a breeding herd of around 1.4 million animals. Beef production was 904 thousand tonnes (carcase weight and including cull dairy cows) (Defra, 2024b).

The wtp calculated above equates to £197 per animal slaughtered and used for breeding, or £0.74/kg of beef produced (carcase weight) per year.

To put the above analysis into context, current retail (supermarket) prices of beef range greatly from around £5 for mince to £35 (AHDB, 2024), although prices for prime cuts can be considerably more (e.g. Waitrose No1 British centre cut fillet beef steak at £60/kg). £0.75/kg is well within the variation of beef prices where premia are paid for higher quality.

Note that the quantity of beef sold retail would be somewhat less than the carcase weight as a percentage of the weight will be lost as a result of the butchery process.

### **Laying hens case study**

The expert panel scored three laying hen scenarios (i) colony cage production with welfare score 32.2 (ii) barn production with score 43.8 and (iii) free-range production with score 51.3.

Using the wtp value tables, an increase in colony cage production welfare score of 32 to barn score 44 (rounded) is valued at £4.44 per one-point increment in welfare score which has a cumulative value of  $12 \times £4.44 = £53.28$ /household/year.

If the entire UK population of laying hens were kept in colony cages the aggregate wtp to move them all to barn production systems would be £53.28 multiplied by 28.4M UK households = £1,513M/yr.

However, 20.8% of eggs are produced from hens kept in colony cages (Defra, 2024c), so the above aggregate value of £1,513M/yr is multiplied by 0.208 = £315M/yr.

An increase in barn egg production score of 44 to free-range egg production at score 51 has a cumulative wtp value of  $7 \times £4.44 = £31.08$ /household/yr, which multiplied by 28.4M UK households = £883M. However, only around 6.6% of eggs are barn eggs (Defra, 2024c), so the above aggregate value is multiplied by 0.066 = £58M.

If all current colony cage egg production switched to free -range, then the value of this based on aggregate wtp across all UK households would be (£1,513M + £883M) multiplied by 0.208 = £498M/yr. The number of eggs produced in the UK, in the four quarters Q4 2023 and Q1, Q2 and Q3 in 2024, was 11,942M of which around 20.8% or 2,484M eggs were from hens in colony cages (Defra, 2024c). The aggregate wtp to increase the welfare of caged hens to that of free-range hens (score 32 to 51) is the equivalent of £0.20 per egg on the current price of cage eggs.

To put the above analysis into context, it is relevant to consider egg prices paid by consumers in the UK. Current retail prices of fresh shell eggs vary greatly depending on egg size, pack size, method of



production, origin, special characteristics (such as breed of hen) and retailer. Retail (supermarket) prices of caged eggs can be found for 13p/egg (Iceland mixed-sized eggs). Free-range egg prices vary from around £0.23/egg (Aldi pack of 12 large British free-range eggs) to £0.66/egg (Waitrose Duchy organic very large eggs pack of six). If the £0.20 wtp per egg estimated above was added to the retail price of caged eggs of £0.13/egg this would give a price of £0.36/egg which is well within the price range for free-range eggs.

Retail prices used in the above policy case study analyses were accessed between 12/11/2024 – 19/11/2024 from retailer websites.

## **Discussion**

A protocol for the assessment of the farm animal welfare impacts of different policy scenarios, on a 0-100 scale, has been developed using an expert panel. In conjunction, a method for estimating the values that UK households place on improvements to the welfare score has been developed, tested and implemented by means of a UK-wide survey of main household food shoppers to elicit their willingness to pay to increase the welfare scores of the six different types of farm animals considered in the study. These methods have been applied to the analysis of six policy case studies.

### **Robustness of the method**

Considerable care has been taken throughout the study to ensure a method and results that are as robust as possible. In particular, the project has:

1. Used best practice methods for both the expert panel welfare score elicitation and for the valuation survey and estimation of willingness to pay values.
2. Built on the experience and findings of previous studies (e.g. that of Bennett et al, 2012).
3. Undertaken extensive testing of methods as part of the development process to ensure that the methods used are fit for purpose and function as required in terms of the production of outputs. (For example, the findings were very similar to each other in the exploratory surveys used to test the valuation method and very similar to those in the main survey).
4. Undertaken qualitative research to help understand consumer thinking and consumer responses to the survey questions (especially to the choice questions) which was used to inform the design of the methods used.
5. Elicited and undertaken analysis of reasoning behind the expert panel scores and behind people's choice selection in the consumer survey to check that the reasoning is both rational and well considered.
6. 'Sense-checked' valuation results with values (i.e. prices) observed in real markets and with consumers' expenditures on the different animal food products.
7. Had continuous internal challenge from Defra and from the project Steering Group regarding methods and results.
8. Provided full details of methods and results for open scrutiny.

## **Limitations**

The limitations identified below are those commonly associated with the use of expert panels, welfare assessment, stated preference valuation and similar studies.

### *Expert panels*

The elicitation of expert judgement using expert panels is widely used to think about complex issues and inform important decisions. Expert panels are particularly helpful where data are lacking. However, the reliability of expert judgement is always sensitive to which experts are involved, the task presented to them and how questions are asked. A number of biases can affect the quality of expert judgements and expert panels. These include those resulting from poorly specified questions and tasks, the difficulty experts may have in providing their judgements in quantitative terms, 'groupthink' (a failure to consider alternative perspectives), a lack of representativeness, partial facilitation/ leadership of the panel, anchoring and contextual biases.

However, the use of an expert panel structured elicitation protocol, as used in this project, can help mitigate various sources of bias and reduce their influence, improve the quality of expert judgements and enhance the accuracy, transparency and defensibility of the judgements elicited from the panel. A diverse group of independent animal scientists was recruited for the panel with different perspectives regarding animal welfare and its assessment.

### *Animal welfare assessment*

There is no single accepted method for assessing animal welfare. Different methods have their advantages and limitations.

However, the Five Domains of Animal Welfare is a widely accepted framework for considering animal welfare and for its assessment (Mellor, 2017). In addition, in the context of farm animal welfare, the Welfare Quality protocols have been widely used and adapted to facilitate on-farm welfare assessment. This project has used a method for welfare assessment that combines both the Five Domains framework and the basis of the Welfare Quality welfare Criteria and Principles to develop a farm animal welfare assessment protocol suited to the elicitation of expert judgement and to the holistic consideration of the welfare of diverse farm animal populations.

### *Economic valuation and people's preferences*

There are only imperfect and imprecise ways to gauge people's preferences. People do not express their preferences perfectly either in stated preference studies or in actual markets because they do not have full information and cannot easily process the information available.

People's preferences are often poorly formed and constantly changing. They rely heavily on the context of decision making. When presented with a new choice decision either in the marketplace (e.g. a new food product in store) or in a stated preference study (e.g. involving animal welfare scores), people may have to construct/re-construct their preferences very quickly as they may not be able to rely on habitual thinking and behaviour.

Thus, any stated preference survey and resulting willingness to pay measures are imprecise estimates of people's preferences and of their actual willingness to pay at any one time (which depends on the context at the time of decision making).

## Conclusion

The project has developed a welfare assessment protocol and implemented it using an expert panel of independent animal scientists. The panel produced well-considered and well-documented scores for a series of example policy scenarios covering six major farm production systems. The welfare scores generated appear reasonable in relation to each other and in relation to the welfare characteristics of the policy scenarios that were assessed. Panel members provided documented rationales and thinking behind their scores panel in terms of the welfare considerations involved which help to explain and justify their judgements and assessments.

The valuation survey elicited welfare score willingness-to-pay values from respondents across the six farm species and for a range of welfare score levels. A stratified sample of over 3,000 respondents was obtained ensuring that the survey was representative of the UK population. The survey method was tested and improved as a result of four exploratory surveys, six focus groups and eight in-depth interviews which provided insights on how well respondents completed the survey questionnaire and their thinking behind their responses. Respondents generally understood the information and task presented to them and had confidence in their responses. People's willingness to pay to increase the welfare score for each species was estimated using appropriate econometric methods which produced estimates that were within the bounds of people's willingness to pay amounts indicated by current retail consumer prices. Tables of valuations of increases to the welfare scores of the six farm animal types have been produced that are largely transferable across policies and can be used to provide valuations of a large range of policy scenarios affecting any of the six animal types.

There is considerable potential for the tools developed and tested in this project to be rolled out to provide economic assessments of policy changes that impact on farm animal welfare. Moreover, in principle, the method could be modified and applied to different animal types.

## References

AHDB, 2024 [ahdb.org.uk/supermarket-red-meat-prices](https://ahdb.org.uk/supermarket-red-meat-prices) (Accessed 14/11/2024)

Balcombe, K.G., Bradley, D., Fraser, I.M. and Hussein, M. (2016). Consumer Preferences Regarding Country of Origin Labelling for Multiple Meat Products. *Food Policy*, 64: 49-62.

Balcombe, K.G., Fraser, I.M. and Di Falco, S. (2010). Traffic Lights and Food Choice: A Choice Experiment Examining the Relationship Between Food Labels and Price. *Food Policy*. 35(3): 211-220.

Bank of England, 2024 Inflation calculator [Inflation calculator | Bank of England](https://www.bankofengland.co.uk/inflation-calculator) (accessed 19/11/2024)

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

Burgman, M.A. (2015) *Trusting judgements: How to get the best out of experts*. Cambridge University Press, Cambridge, UK. 203 pages.

Butler, A.J., Thomas, M.K. and Pintar, K.D.M. (2015) Systematic review of expert elicitation methods as a tool for source attribution of enteric illness. *Foodborne Pathogens and Disease*. Vol 12(5), 367-378.

Defra (2024a) Usage of milk by dairies in the UK statistics: August 2024. <https://www.gov.uk/government/statistics/historical-national-statistics-notice-on-milk-utilisation-by-dairies-2024> (Accessed 15/11/2024).

Defra (2024b) [Livestock populations in the United Kingdom at 1 June 2023 - GOV.UK](#) (Accessed 15/11/2024).

Defra (2024c) Quarterly UK statistics about eggs – statistics notice (data to Q3 2024). <https://www.gov.uk/government/statistics/egg-statistics> (accessed 19/11/2024).

Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F. and Wintle, B.C. (2018) A practical guide to structured expert elicitation using the IDEA protocol. *Methods in Ecology and Evolution*, Vol 9, 169-180.

HM Treasury Green Book <https://www.gov.uk/government/collections/the-green-book-and-accompanying-guidance-and-documents>. (Accessed 22 January 2025).

[Kehlbacher, A.](#), [Bennett, R.](#) and [Balcombe, K.](#) (2012) *Measuring the consumer benefits of improving farm animal welfare to inform welfare labelling*. *Food Policy*, 37 (6). pp. 627-633. ISSN 0306-9192 doi: <https://doi.org/10.1016/j.foodpol.2012.07.002>

[Kehlbacher, A.](#), [Balcombe, K.](#) and [Bennett, R.](#) (2013) *Stated attribute non-attendance in successive choice experiments*. *Journal of Agricultural Economics*, 64 (3). pp. 693-706. ISSN 0021-857X doi: <https://doi.org/10.1111/1477-9552.12021>

Mellor, D. J. (2017) Operational Details of the Five Domains Model and Its Key Applications to the Assessment and Management of Animal Welfare. *Animals* Aug 9;7(8):60.)

Welfare Quality Network <http://www.welfarequality.net> (accessed 19/11/2024)