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

Final Report

Richard Bennett, Kelvin Balcombe and Heather Maggs

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

This research was funded by the UK Department for Environment, Food & Rural Affairs

https://doi.org/10.48683/1926.00123483.

January 2025

Contents

Annexes

Summary

There is a requirement to assess the costs and benefits associated with government policy and of policy changes (HM Treasury Green Book, 2025). However, 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 are transferrable between policies and policy appraisals. A method has been developed for the assessment of the animal welfare impacts of policy designed to improve the welfare of farm animals using a 0-100 scale. This is combined with a choice experiment stated preference survey which elicited food shoppers' willingness to pay (wtp) for increases to the welfare scores of six different farm animal types (laying hens, dairy cows, indoor pigs, sheep, broiler chickens and beef cattle). The welfare score provides a common link between the welfare assessment and the economic valuation.

A structured elicitation protocol was developed for the assessment of changes to farm animal welfare as a result of a policy change. An expert panel of 13 independent animal scientists provided assessments of welfare for a range of policy scenarios on the 0-100 scale. Zero denotes extreme suffering of animals and 100 denotes the highest achievable welfare possible. Twelve welfare Criteria were scored by the panel including consideration of the mental states of animals and their whole life experience from birth to slaughter, with detailed reasoning behind their scoring.

Valuation of the animal welfare impacts in terms of people's wtp for increments in the 0-100 score was elicited via an online choice experiment survey, from a representative sample of over 3,000 main household food shoppers in the UK. The method was extensively tested using four exploratory surveys, six focus groups and eight in-depth interviews. The survey contained choice questions showing different combinations of animal welfare scores and increases in weekly household food bill. People's responses to the choice questions were used to estimate their wtp (£s/household/year) for successive one-point increments in the animal welfare score for each farm animal type, using a Hierarchical Bayesian Logit statistical model. The table below provides examples of cumulative wtp values at different points of the welfare scale from the status quo starting points used in the survey. Marginal wtp diminishes as welfare score increases. Full valuation tables are provided in the report.

	SQ	SQ+1	SQ+5	SQ+10	SQ+15	SQ+20	SQ+30
Laying hens	50	4.44	21.40	40.80	58.21	73.62	98.43
Dairy cattle	40	5.18	25.15	48.44	69.87	89.43	122.95
Indoor pigs	40	4.27	20.74	39.93	57.59	73.72	101.35
Sheep	50	4.04	19.47	37.13	52.96	66.98	89.57
Broilers	40	4.56	22.14	42.65	61.51	78.72	108.24
Beef cattle	60	4.40	20.94	39.25	54.91	67.94	86.06

For example, if a policy is assessed to increase the welfare score of broilers from 40 to 45, the wtp associated with this from the table is £22.14/household/yr which multiplied by the number of UK households (28.4M) gives a total benefit of £628.8M/yr. The wtp welfare score valuations are transferable across different policies which impact on the animal types considered and should be usable for some years.

The merits and limitations of the method and outputs produced are presented. The major merit is that the method is robust, with credible outputs which are transferable across policies and which should be valid to use for some years. There is considerable potential for the method to be rolled out to provide economic assessments of a range of policies and policy changes that impact on farm animal welfare.

Executive Summary

There is a requirement to assess the costs and benefits associated with government policy and policy changes (HM Treasury Green Book, 2025). However, 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 are transferrable between policies and policy appraisals.

The aim of the project was to develop a method for the assessment of the animal welfare impacts of policy designed to improve the welfare of farm animals on a 0-100 scale and to use a choice experiment stated preference survey to elicit consumers' willingness to pay valuations for increases to the welfare scores of six different farm animal types. The welfare score provides a link between the welfare assessment and the economic valuation parts of the method.

An expert panel, structured elicitation protocol was developed for the assessment of changes to farm animal welfare as a result of a policy change. The protocol used the Five Domains of Animal Welfare model and the 12 welfare Criteria and four Principles from Welfare Quality protocols but these were expanded to include consideration of the mental states of animals and their whole life from birth to slaughter. For a given policy scenario, the expert panel (of 13 independent animal scientists) used the welfare assessment protocol to score each of the welfare Criteria on the 0-100 scale where 0 denotes extreme suffering of animals and 100 denotes the highest achievable welfare possible. Welfare Criteria were then combined into four welfare Principles and the four Principles into a single welfare score using Choquet Integral weightings elicited from the panel (which generally gave greater weighting to low welfare scores).

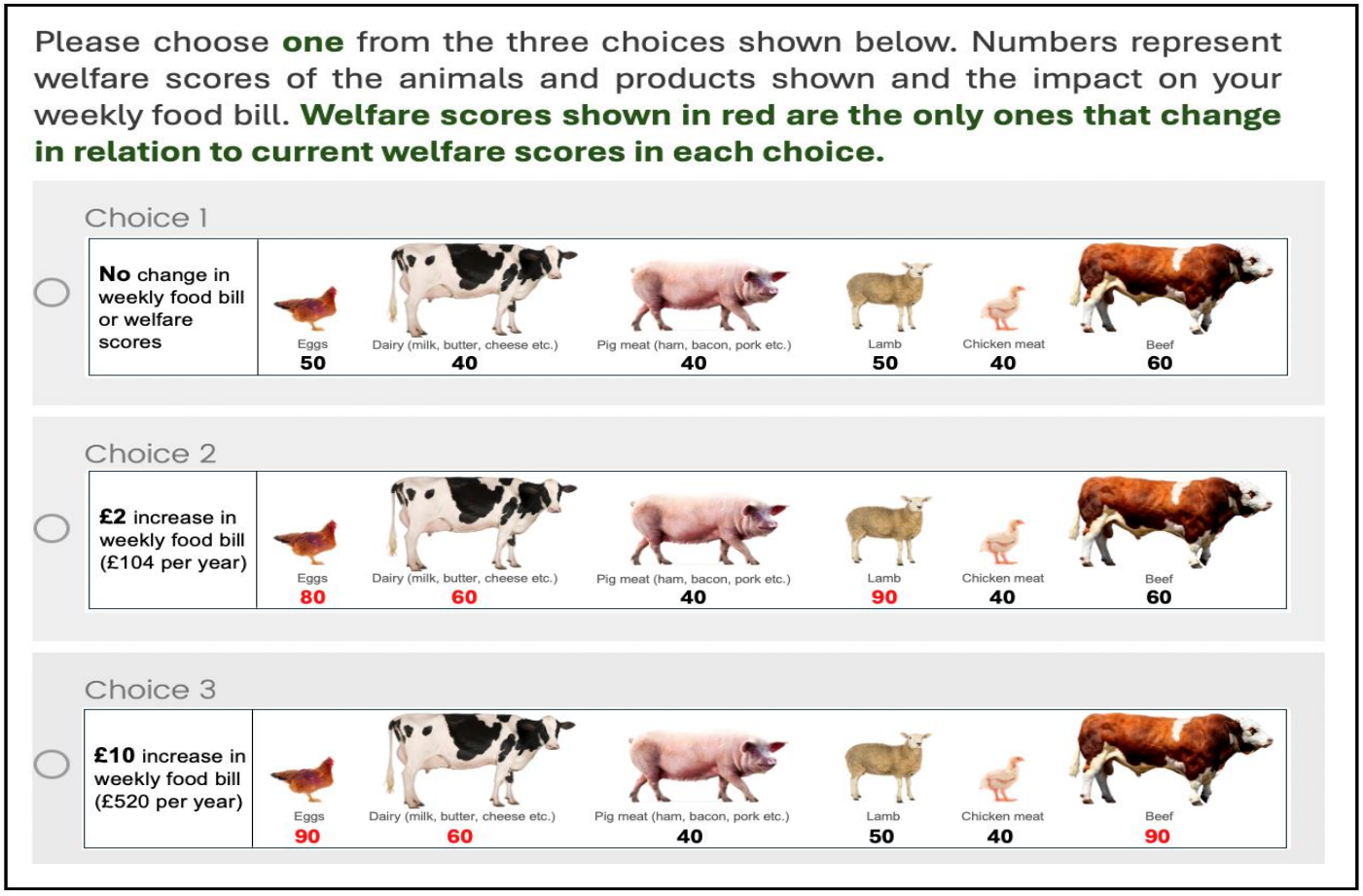
Six policy case studies were selected (each with two or three policy scenarios) to include a range of topical hypothetical policies to test the welfare assessment method with coverage across the six farm animal types. Expert panel members provided reasoning behind their scoring for each policy scenario. The final single panel scores for the policy scenarios are shown in Table 1.

Table 1. Policy scenario scores

Species	Production System	Choquet Integral Score		Species	Production System	Choquet Integral Score
Laying Hens	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
Sheep production	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				
Beef production	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

Economic valuation of the animal welfare impacts in terms of people's willingness to pay for increments in the 0-100 score was elicited via an online choice experiment questionnaire survey, from a stratified sample of over 3,000 consumers in the UK who were the main food shoppers for their household. Extensive testing of the method was undertaken including four exploratory surveys, six focus groups and eight in-depth interviews. Respondents to the survey were presented with 12 choice questions (sets) each with three options from which they had to choose their preferred option, with six different groups of choice sets randomly allocated to respondents. 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 and consumption is done at the household level. An example choice question is shown in Figure 1. People's responses to the choice questions were used to estimate their willingness to pay, using a Hierarchical Bayesian Logit statistical model, for successive one-point increments in the animal welfare score for each of the six farmed animal types (laying hens, dairy cows, indoor pigs, sheep, broiler chickens and beef cattle).

Figure 1. Example choice set



Eighty-nine percent of respondents understood the information presented to them in the questionnaire and 86% had confidence in their responses to the choice questions.

Table 2 shows a sample of results from the estimation of marginal willingness to pay for successive one-point increases in the welfare score from the status quo values used in the survey. The results show declining marginal wtp values (£/yr) for a one-point increase in welfare score as the score increases at different points from the status quo (SQ) welfare levels for each animal type in the UK. Table 3 shows associated cumulative wtp values.

Table 2. Marginal willingness to pay values (£'s per household per year) for one point increase in welfare score

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

Table 3. Cumulative willingness to pay values (£'s per household per year)

	SQ	SQ+1	SQ+5	SQ+10	SQ+15	SQ+20	SQ+30
Laying hens	50	**4.44**	**21.40**	**40.80**	**58.21**	**73.62**	**98.43**
Dairy cattle	40	**5.18**	**25.15**	**48.44**	**69.87**	**89.43**	**122.95**
Indoor pigs	40	**4.27**	**20.74**	**39.93**	**57.59**	**73.72**	**101.35**
Sheep	50	**4.04**	**19.47**	**37.13**	**52.96**	**66.98**	**89.57**
Broilers	40	**4.56**	**22.14**	**42.65**	**61.51**	**78.72**	**108.24**
Beef cattle	60	**4.40**	**20.94**	**39.25**	**54.91**	**67.94**	**86.06**

A simple example helps to show how the values can be used. Suppose a policy is likely to improve the welfare of broiler chickens from an assessed welfare score of 40 to a score of 45. The benefit of this policy in terms of willingness to pay could be read from the cumulative wtp valuation table as £22.14/household/yr. To obtain the total UK benefit this amount can be multiplied by the number of UK households (around 28.4M) to give a total estimated benefit of the policy change of £628.8M/yr. Full wtp valuation tables are provided in the main report. Table 4 shows benefit calculations for each of the policy scenarios considered.

Table 4. Benefit calculations

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^2(c) to 30kg/m^2	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^2. (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.

The merits and limitations of the method and outputs produced are presented. The major merit is that the method is robust, with credible outputs which are transferable across policies and which should be valid to use for some years. There is considerable potential for the tools developed and tested in this project to be rolled out to provide economic assessments of a range of policies and policy changes that impact on farm animal welfare.

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.

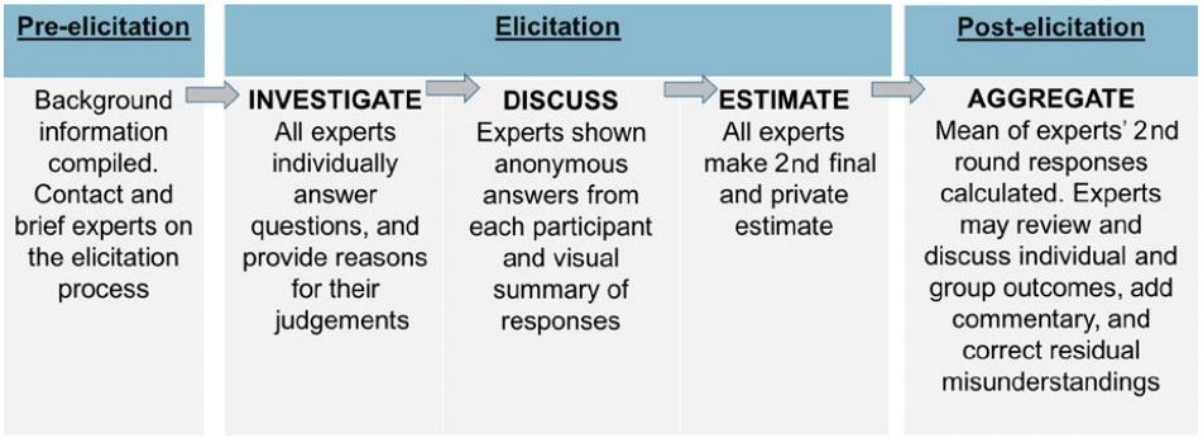
7. Production and sharing of information with the expert panel in relation to each of the policy case studies and scenarios.

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 mainly 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. An example score sheet, which is populated with the 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 were 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 were 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 were asked to provide ‘first round’ scores for each of the welfare Criteria in relation to each of the policy scenarios being considered. They also supplied 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 returned their scores to the facilitators and the scores were collated and fed back to the panel so that members were able to see each other’s anonymized scores prior to meeting to discuss them. Panel members were 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 contained the panel’s score weightings (see below for further details regarding this Choquet Integral function).

Table 1. Example extract of an expert panel welfare assessment score sheet with a panel member's 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**	0 – Lowest level of welfare … Highest level of welfare = 100	**1st round score**	**2nd round score**	**Comments Round 1**
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 preventitive 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 met 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 was chaired by a facilitator supported by the project lead and a person responsible for collating the scores. The facilitator led discussion regarding the panel's scores, going through each of the Criteria one by one. Prior to this, panel members were reminded that they will have the opportunity to revise their scores if they wished, 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 were 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 were 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 were reviewed and discussed by the panel in the meeting. The panel considered whether the final aggregate scores adequately represented 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 that 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 as part of the Welfare Quality project showed that the weightings were not substantially different between animal types.

The method for calculating the Choquet weightings is described in the detailed report version in Annex 16. Resultant Choquet Integral weightings from the expert panel 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 policy scenarios 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 in relation to the numbers and types of animals affected, the duration and magnitude of impact for animals affected and the nature of the production system.

The six policy case studies 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/m2 to 30kg/m2) (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 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.

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

Species	Production System	Choquet Integral Score		Species	Production System	Choquet Integral Score
Laying Hens	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
Sheep production	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				
Beef production	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

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. 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 was directed at citizens who are the main food shoppers in their household in the UK. The survey was 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 of the exploratory research and its findings are provided in the detailed report version in Annex 16.

Final survey design

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

1. Some socio-economic and demographic questions to allow quota sampling of the population such as whether respondents 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 the 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 animal product consumption patterns but also focused participant's minds on 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 (i.e. the welfare score) 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 in 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 was an open text box that asked respondents to briefly state their reasoning. The second question asks whether they paid attention to specific 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.

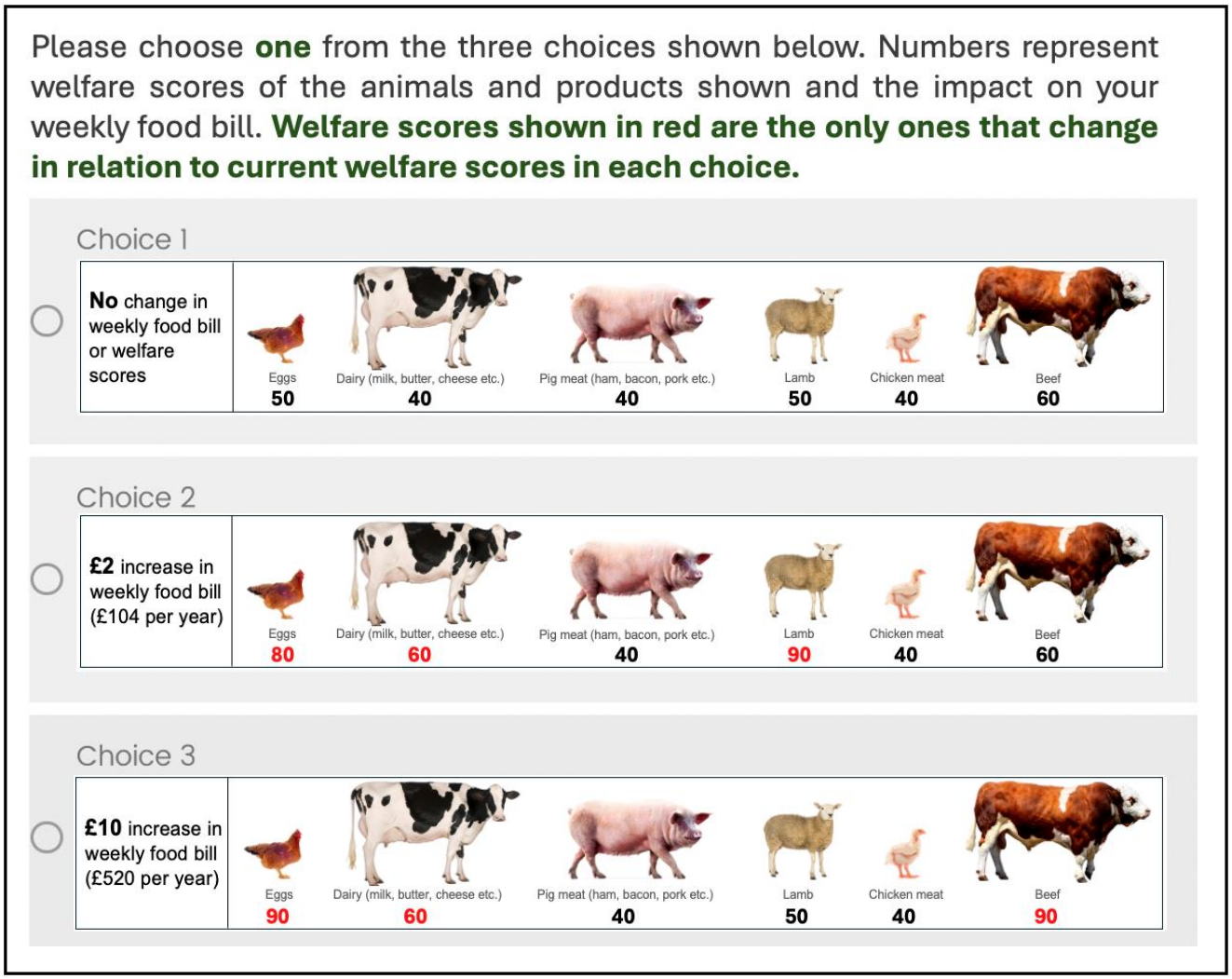
Figure 2 is an example of a choice set with 12 similar choice sets being presented to each individual respondent in the questionnaire.

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 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.

Figure 2. An example of a choice set in the survey questionnaire



None of the options showed a reduction in welfare score relative to the status quo level for each animal type (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.

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 3rd 2024 and completed on October 15th 2024 with a total of 3,013 completed questionnaires obtained. Annex 14 shows the quota levels in the survey.

To analyse the discrete choice experiment data, a Hierarchical Bayesian Logit (HBL) was employed (Balcombe et al. 2016). Two models were explored, a basic one more commonly used for such analyses and an extended version. Details of the methods of analysis are provided in the detailed report version in Annex 16. Importantly, the extended HBL model allowed us to estimate marginal willingness to pay (wtp) amounts for successive one-point increases in the welfare score for each animal type (starting from the status quo score values used in the survey) across the range of animal welfare scores.

Results

Full details of the nature of the survey sample population (e.g. age, income, education, ethnicity etc) with Tables are provided in the detailed report version in Annex 16. It should be noted that a number of characteristics of the sample were largely determined by the quota requirements of the survey used to ensure that a sample that was representative of the UK population was obtained.

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). Sixty-five percent of respondents were female and 35% male.

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

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.

Table 3 below shows household expenditure on food each week.

Table 3 Household expenditure on food (£/week)

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

Table 4 shows household expenditure on specific animal products.

Table 4 Household expenditure on the main animal food products (£s/week)

Weekly household expenditure	% of households					
	Chicken	Beef	Lamb	Pigmeat	Dairy	Eggs
£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

Sixty-six percent of respondents stated that concerns about animal welfare influenced their purchasing decisions. 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% trusted that farms would be properly monitored.

People's willingness to pay to increase animal welfare scores

Results from using the extended Hierarchical Bayesian Logit (HBL) model for analysing the choice experiment data are shown in Table 5.

Mean is the mean willingness to pay (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 an average wtp of £5.18/household/year to increase the welfare of dairy cattle in the UK by one welfare score point from the status quo value.

Table 5. Extended model wtp results

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

Distribution of wtp estimation across the iterations 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 0.256. 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.

A range of wtp values can be seen across the animal types. 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).

Table 6 below shows additional results from the estimation of marginal willingness to pay (£s/household/yr) 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 6. Marginal wtp (£s/household/yr) 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
Hens laying	50	4.44	4.04	3.64	3.24	2.84	2.04
Dairy	40	5.18	4.81	4.43	4.06	3.69	2.94
Pigs	40	4.27	3.96	3.66	3.35	3.04	2.43
Sheep	50	4.04	3.68	3.31	2.95	2.59	1.86
Broilers	40	4.56	4.23	3.90	3.58	3.25	2.59
Beef cattle	60	4.40	3.87	3.34	2.82	2.29	1.23

The table shows that people have a higher marginal wtp at lower levels of welfare score and lower levels at higher scores. For example, for laying hens, marginal wtp falls to a value of £4.04 per year for a one-point increase from welfare score 54 to 55 (SQ+5.00), £3.64 for a one-point increase from 59 to 60, £3.24 for 64 to 65, £2.84 for 69 to 70 and £2.04 for 79 to 80.

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 6.

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

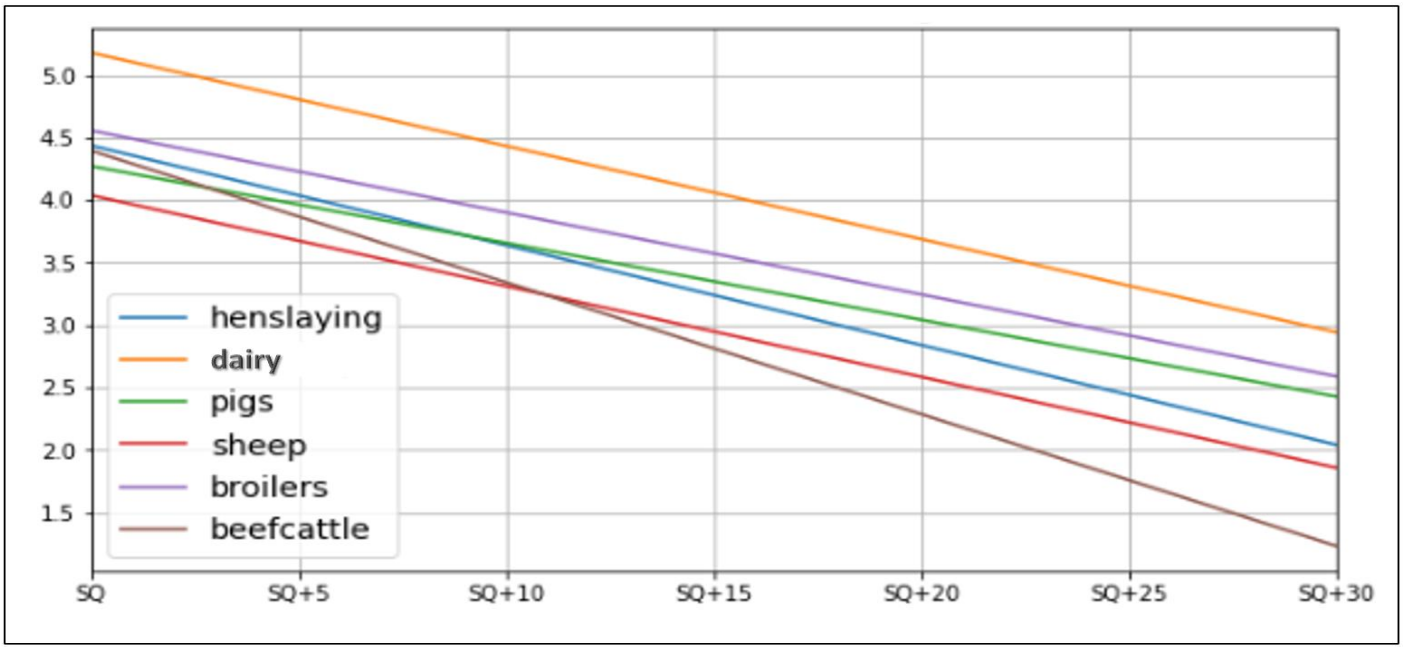


Table 7 shows cumulative marginal wtp values, again as snapshots at different points further from the status quo levels for each animal type.

Table 7. Cumulative wtp (£s/household/yr) at various points from the status quo (SQ)

	SQ	SQ+1	SQ+5	SQ+10	SQ+15	SQ+20	SQ+30
Laying hens	50	**4.44**	**21.40**	**40.80**	**58.21**	**73.62**	**98.43**
Dairy cattle	40	**5.18**	**25.15**	**48.44**	**69.87**	**89.43**	**122.95**
Indoor pigs	40	**4.27**	**20.74**	**39.93**	**57.59**	**73.72**	**101.35**
Sheep	50	**4.04**	**19.47**	**37.13**	**52.96**	**66.98**	**89.57**
Broilers	40	**4.56**	**22.14**	**42.65**	**61.51**	**78.72**	**108.24**
Beef cattle	60	**4.40**	**20.94**	**39.25**	**54.91**	**67.94**	**86.06**

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. These tables are provided in Annex 15. They show marginal wtp values for each one-point increment in welfare score, together with these marginal values aggregated into cumulative values of the one-point increments. These are referred to as the 'wtp valuation tables'.

A simple example, with reference to Table 7, helps to show how the wtp valuation tables can be used. Suppose a policy is likely to improve the welfare of broiler chickens from an assessed welfare score of 40 (for a particular policy scenario) to a welfare score of 45 (for a specified changed policy scenario). The benefit of this policy in terms of willingness to pay could be read from the cumulative wtp valuation table as £22.14/household/yr. This amount can be multiplied by the number of UK households (approximately 28.4M) which would give a total estimated UK benefit of the policy change of £628.8M /yr.

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

Benefit calculations for each of the policy scenarios considered are shown in Table 8. A full description of the calculations and reasoning involved in each of the policy scenario benefit calculations are presented in the extended detailed report of Annex 16.

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.

Table 8. Benefit calculations for each policy scenario[(a)]

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^2(c) to 30kg/m^2	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^2. (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.

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 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 and of results

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 of the method and results

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 (Accessed 14/112024)

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 (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-notices-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 15/11/204).

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)