Is animal welfare better on smaller dairy farms? Evidence from 3,085 dairy farms in Germany

T. Lindena[1]* and S. Hess[2]
[1]Johann Heinrich von Thünen Institute, Institute of Farm Economics, Braunschweig, Lower Saxony, 38116, Germany
[2]Department of Agricultural Markets, University of Hohenheim, Stuttgart, Baden-Württemberg, 70599, Germany

ABSTRACT

The structural change toward larger (dairy) farms is often criticized because it supposedly has a negative effect on farm animal welfare. We investigated this criticism using cross-sectional survey data from 3,085 German dairy farms. Even though our sample was a convenience sample, it closely resembled the diverse structures of dairy farming in Germany and covered a wide range of dairy farm sizes (7 to 2,900 cows per farm, mean 122). We developed an animal welfare index (AWI) in close consultation with experts along the dairy value chain (e.g., farm animal welfare scientists, farmers, dairy representatives). Regression results showed that larger farms tended to achieve a better AWI than smaller farms in our data set. However, the effect size was small. Nevertheless, in contrast to the widespread assumption in public discussion, larger dairy herds are not necessarily associated with poorer animal welfare. In all herd size classes, we found a large variation of AWI between herds. Although this study focused on the effect of herd size, it is not the only factor affecting animal welfare levels on individual farms. Other variables that we included in the regression to describe the AWI indicate that the knowledge and skills of the farm manager and the amount of time that farms can devote to animals have a positive effect on the AWI. However, as with herd size, the effect size of other explanatory variables was small in absolute terms.
Key words: farm animal welfare, herd size, dairy sustainability, dairy cow

INTRODUCTION

The European agricultural sector is undergoing continuous structural change toward increasing concentration of production in fewer, larger, and more specialized farms (e.g., Zimmermann and Heckelei, 2012; European Commission, 2013). For instance, in Germany, the fourth largest milk producer in the world (Hemme, 2020), the mean number of dairy cows per farm increased from 31 cows per farm in 1999 to 70 cows per farm in 2021, an increase of more than 100% in the last 2 decades. In 2021, 1 in 5 farms kept more than 100 dairy cows. Thus, 57% of dairy cows in Germany were kept on farms with more than 100 dairy cows compared with 21% in 1999 (Forstner and Nieberg, 2019; Destatis, 2021). Overall, it is assumed that the structural change will continue. The main driving forces include not only scientific and technological developments and the need to increase total factor productivity but also the additional requirements for livestock farming (such as food safety, environmental requirements, and animal welfare), which have largely a fixed cost character for farms and have steadily increased in recent years (e.g., Sauer and Latacz-Lohmann, 2015; Forstner and Nieberg, 2019).

At the same time, the issue of animal welfare in agriculture is receiving increasing attention from European citizens, who want more information on the conditions in which farm animals are kept (European Commission, 2016). Against this background, public and private animal welfare certification or regulation schemes discuss issues such as housing system, outdoor access for animals, and space and farm size restrictions as possible criteria to ensure animal welfare and maintain consumer confidence. The label program "For more animal welfare" of the German Animal Welfare Association, for example, sets a maximum farm size of 600 cows per farm for participating dairy farms (Deutscher Tierschutzbund e.V., 2021). Most likely, farm size restrictions are introduced because it is often assumed, in public discussion, that there is a negative link between farm or herd size and animal welfare (e.g., Busch et al., 2018; Pfeiffer et al., 2021). Farmers, on the other hand, seem to see no relationship between farm size and animal welfare (e.g., Vanhonacker et al., 2008; Sørensen and Fraser, 2010). Recent scientific research reveals inconclusive results regarding a relationship between animal welfare and farm size.

Animal welfare is defined as a multidimensional concept that encompasses basic animal health and functioning, natural living, and affective state (Fraser, 2008). Due to the multidimensional character of animal welfare, a holistic approach that simultaneously considers several animal welfare indicators is crucial for investigating the effect of herd size on animal welfare (Robbins et al., 2016). Very few studies have taken this type of comprehensive approach, and most of them focused on pig and poultry farming. Some of these studies revealed a higher level of animal welfare on larger farms (Wellbrock et al., 2009), others found negative relationships between animal welfare and farm size (Mazurek et al., 2010), and some do not mention any relationship (Stott et al., 2012). Other studies have assessed animal welfare on the basis of violations of animal care regulations and found either a positive relationship (i.e., fewer violations on large farms; Hess et al., 2014) or no relationship with farm size (Czekaj et al., 2013; Andrade and Anneberg, 2014; Otten et al., 2014).

In dairy farming, Gieseke et al. (2018) used the Welfare Quality animal welfare indicators (Welfare Quality Consortium, 2009) to investigate the effect of herd size on animal welfare in 80 German dairy farms of above-average size. Their findings did not indicate a linear relationship between herd size and overall welfare score. Robbins et al. (2016) examined more than 150 publications considering the relationship between farm size and single animal welfare indicators (e.g., lameness, udder health, or mortality) with a focus on dairy farms. They also found little evidence of any simple relationship, negative or positive, between farm size and animal welfare: "Instead the evidence suggests that larger farms provide some opportunities to improve animal welfare but may also create welfare risks" (Robbins et al., 2016, page 1). Beggs et al. (2015, 2019) emphasize this point: their analyses of Australian pasture-based dairy farms show that some welfare risks increase with herd size but the use of avoidance strategies in this context also increases. Robbins et al. (2016) conclude their comprehensive review by pointing out that, in addition to a multidimensional approach (with several animal welfare indicators), a more sophisticated analysis that simultaneously incorporates several intervening variables could be more elucidating in the debate on animal welfare and herd size, but this approach has rarely been investigated to date.

Against this background, the objective of the present study is to empirically examine the relationships among animal welfare (combined to an index from various animal welfare indicators), herd size, and other explanatory variables on a large number of German dairy farms, with a focus on investigating the effect of herd size. For this purpose, we summarize existing hypotheses on the relationship between animal welfare and herd size. Positive, negative, and curvilinear relationships between animal welfare and herd size are theoretically conceivable. Alternatively, these two variables may not exhibit any statistically significant relationship at all. We empirically tested these hypotheses using farm survey data from Germany. To achieve our goal, we developed an animal welfare index (**AWI**) in close consultation with experts along the dairy value chain.

MATERIALS AND METHODS

Data from the German Dairy Sustainability Tool

The present study used cross-sectional data collected as part of a nationwide dairy sustainability project (Dairy Sustainability Tool, **DST**) involving more than 30 German dairies (more than a quarter of all dairies in Germany) and their supplying farmers. It has been contractually agreed with the participating dairies that they regulate data protection and a data use agreement with their supplying dairy farmers (who complete the questionnaire) in accordance with legal requirements. The questionnaire, which was distributed to participating farmers via the dairies, covers all dimensions of sustainability, including economic, ecological, and social issues, as well as aspects of animal welfare. A central database was implemented for the DST survey. There were 3 ways to add data to the database (i.e., to complete the questionnaire): (1) a dairy farmer entered the data into the central database via a web-based questionnaire; (2) a third person, usually an independent person who collects data for the dairies for other audits on the farms (e.g., quality audits) came to the farm and completed the questionnaire prepared by the farmer together with the farmer, using a tablet computer, in which case the auditor was directly available for any queries; or (3) the farmer filled out the questionnaire on paper and dairy staff then entered the data into the central database. Dairy staff contacted farmers during the entry procedure if there were any discrepancies. Regardless of the type of data collection, initial plausibility checks were programmed into the DST survey. In the event of implausible data entries, the person entering the data would receive a warning or error message directly and was instructed to check the value and correct it if necessary (e.g., if the number of lactating cows entered at the beginning and that entered later for different barns do not match). In addition, upper and lower limits were set in the online input screen that represented the limits of the possible range (e.g.,

maximum values for the average lifetime performance of culled cows that can be entered).

The underlying data were collected between June 2017 and April 2020. The final data set comprised 3,085 farms with 376,415 dairy cows, corresponding to 5.6% of dairy farms and 9.8% of dairy cows in Germany. For a more detailed description of the data set, especially on herd size, see section "Descriptive Statistics of Covariates."

Because only dairy farmers belonging to dairies that participated in the DST were able to complete the questionnaire, we designated this to be a convenience sample. The DST is, in principle, open to all dairies in Germany, but participation is voluntary. Most participating dairies attempted to conduct a full survey of all their milk-supplying farmers but were primarily dependent on the voluntary participation of their farmers. The dairies therefore intensively promoted farmer participation in the DST; for example, through presentations, newsletters, or personal contact. Some dairies also offered farmers a small compensation for the time spent completing the questionnaire. An additional incentive for farmers to participate was that they receive a farm-specific sustainability benchmark in return. We assume, therefore, that the sample sufficiently covers many farms with no initial interest in the DST, although we cannot rule out the possibility that some dairies and farmers with the least interest in the DST systematically avoided participation in the survey. The mean response rate across all dairies was 62%; one dairy achieved a response rate of 100%. The lowest response rate was 13% and was achieved by a dairy that implemented the topic of sustainability more intensively with farmers for the first time and was met with skepticism from farmers.

Because the issue of animal welfare has increasingly become a focus of public attention, some respondents might have chosen to answer in a socially desirable way. However, before data collection, dairies clearly communicated to their farmers that the goal of the study was an assessment of the status quo, with neither rewards nor penalties for any stated animal welfare level. On the whole, the results of the survey suggest that the farmers answered honestly because, for example, statements that did not comply with legal regulations could be found in the answers. Furthermore, the structure of the questionnaire offered many plausibility cross-checks. If implausible statements emerged during data checks, these were either clarified via the dairies with the respective dairy farmer or deleted from the data set. Overall, we cannot completely exclude the possibility that some answers were leaning toward more socially desirable outcomes, but substantial effort was made to minimize this effect.

Definition and Measurement of Animal Welfare

Animal welfare is dependent on personal and societal values and ideals. Consequently, animal welfare is neither clearly definable nor conclusively measurable by science (Mason and Mendl 1993). Thus far, the definition of Fraser (2008) has been widely accepted: animal welfare is a multidimensional concept that encompasses basic animal health and functioning, natural living, and affective state and must be considered as a multidimensional whole (Botreau et al., 2007; Marchant-Forde, 2015). The "Five Freedoms" (Table 1) published by the Farm Animal Welfare Council (FAWC, 1979) cover the 3 areas from Fraser (2008) and frequently serve as the basis of a holistic animal welfare assessment.

When attempting to measure animal welfare, it is possible to differentiate between so-called resource-, management-, and animal-based indicators (Roesch et al., 2016). The choice of indicators to measure animal welfare depends largely on the purpose of the data collection (e.g., Sørensen and Fraser, 2010). Several animal welfare assessment tools are available, each developed for specific user groups and purposes, such as for self-monitoring by farmers, animal-welfare checks in the scope of advisory services, evaluation systems such as the "Welfare Quality Assessment Protocol" for primarily scientific purposes, and various trade and marketing labels.

Until now, the Welfare Quality protocol has been considered the best available tool for scientifically sound animal welfare assessment. However, due to the direct assessment of animals by an expert on site, the collection of (mainly) animal-based indicators is very time-consuming and therefore not suitable for a broad application on a large number of farms (Robbins et al., 2016; Roesch et al., 2016). Consequently, compromises have to be found to reconcile scientific knowledge and practicability, which includes, in particular, cost-effective and efficient data collection. Against this background, more pragmatic solutions are often used, which results in the use of mainly resource- and management-based indicators (Roesch et al., 2016). Other authors also emphasize, in addition to validity and repeatability, feasibility as the main criterion for applicability in practice (e.g., Scott et al., 2001; Waiblinger et al., 2001). Roesch et al. (2016) point out that the use of resource-, management-, *and* animal-related indicators could provide a more valid representation of animal welfare.

The DST attempts to support as many farms as possible in their development toward increased sustainability and animal welfare. Thus, the tool's main purpose is a holistic farm self-assessment to didactically assess the strengths and weaknesses of a farm and serve as a

Table 1. Selected animal welfare indicators of the Dairy Sustainability Tool and assignment to the 4 categories of the Welfare Quality Protocol and the Five Freedoms

Assignment to the 4 categories of the Welfare Quality Protocol[1]				Assignment to the Five Freedoms[1]: freedom from/to...					Selected animal welfare indicators
Good housing	Good feeding	Good health	Appropriate behavior	Hunger and thirst	Discomfort	Pain, injury, or disease	Express normal behavior	Fear and distress	
X			X		X		X		Freedom of movement for the dairy cows (husbandry system)
X			X		X		X		Access to outdoor climate stimuli for dairy cows; dry cows; young cattle
X		X	X		X	X	X		Type of cubicle and cubicle floor covering; management of lying areas
X		X	X			X	X	X	Ratio: resting areas to dairy cows
	X	X	X	X		X	X		Ratio: feeding areas to dairy cows
	X	X	X	X	X	X	X		Number of drinking troughs; inspection and cleaning of drinking troughs
X					X		X		Existence of cow brushes; facilities for thermoregulation
X			X		X	X	X	X	Existence of calving pens; management of calving pens; existence of areas for sick dairy cows
		X	X		X	X			Proportion of difficult calvings
		X			X				Proportion of dairy cows with ≤100,000 SCC/mL at average for the last 12 mo; way and frequency of maintenance of the milking system
		X				X			Strategy for the use of antibiotics for mastitis
		X	X			X			Control of joint injuries
		X	X			X			Performing cow-individual lameness control
		X				X			Frequency of prophylactic hoof care
	X	X		X	X	X			Metabolism profile of dairy cows: frequency of analysis; share of dairy cows with a fat:protein ratio ≥1.5 or <1 for the last year
		X				X			Dairy cow mortality rate for the last year
	X	X				X			Herd care by external experts
X		X	X		X	X			Care of resting areas for newborn calves
	X	X	X	X	X				Calf rearing management: supply of colostrum; supply of water
		X				X			Disbudding method
		X				X			Calf mortality rate for the last year

[1]On the basis of these assignments we derived the weights for alternative index calculations (see Methods section).

basis for management improvement or strategy development. In addition, it serves as a monitoring scheme for the dairy industry. In this context, the DST focuses primarily on resource- and management-based indicators (Table 2). Several animal-related indicators have, however, been included, such as udder and metabolic health and the rate of calving difficulties. These data are available for most dairy farms due to their participation in national milk yield recording programs (Flint et al., 2016).

Identification and Selection of Animal Welfare Indicators and Their Respective Assessments

The identification and selection of animal welfare indicators and their respective assessments have been carried out in previous projects. In a first step, scientifically based indicators for measuring sustainability, including animal welfare, were compiled following an extensive literature review. Furthermore, a broad range of existing sustainability and animal welfare assessment tools were analyzed in terms of topic and indicator selection. Because international connectivity of the DST is an important goal, we analyzed not only indicator catalogs and assessment tools tailored to the German situation, but also those that exist in other important milk-producing countries. Finally, current requirements of the market partners (industry customers and food retailers) for the dairies were included in the work. In view of the above-mentioned objective of involving as many dairy farmers as possible in the survey, it was important to keep an eye on practicability: the indicators had to be ascertainable in the scope of a written survey with justifiable effort. Based on this, a questionnaire was developed to record selected sustainability indicators including animal welfare in dairy production. This was followed by initial surveys to extensively test the practicability of the indicators in a questionnaire survey (the results of which are presented in detail in Lassen et al., 2014, 2015). The accumulated knowledge and the initial survey experience led to the preselection of indicators (Flint et al., 2016).

In a second step, assessments were developed for the preselection of indicators in the form of a 4-point scale, where "level 3" indicates the optimal outcome in terms of animal welfare and "level 0" represents a threshold. The content of the indicator assessments was based on (1) scientific evidence on the respective indicator, (2) legal regulations, (3) available ratings in existing animal welfare assessment tools, and (4) known distributions of practical data from statistics on individual indicators. For each indicator, a factsheet was prepared with detailed descriptions and an assessment approach (Flint et al., 2016). The assessment categories do not appear

in the questionnaire and were therefore not known to farmers at the time of the survey. The questionnaire was structured in such a way that the farmers selected those qualitative items that reflected the actual situation on their farm during the survey. The assessment categories (4-point scale) were then calculated, often from more than one question. For example, for the indicator "Existence of areas for sick dairy cows," farmers could not simply check one of the assessment categories (e.g., Level 1: Separate area for sick cows available, but for less than 2% of the herd or less than 10 m^2/cow or both; see Supplemental File S1, https://www.openagrar.de/receive/openagrar_mods_00081639) directly. Rather, the assessment category was calculated from a total of 5 questions: a farm has indicated that a rehabilitation area is available (question 1) and exactly 1 (question 2), which has 15 m^2 per cow (question 3). In total, this rehabilitation area is enough for a maximum of 2 cows (question 4). However, the farm indicated at the beginning of the questionnaire that it has a total of 200 cows (question 5). Thus, the farm has a rehabilitation area for only 1% of its herd, and this farm is scored as "level 1." An excerpt from the questionnaire can be found in Appendix 1.

Step 3: Building on this, a large multi-stakeholder workshop was conducted in 2016, with experts along the dairy value chain and scientists. In the workshop, the potential animal welfare indicators and the corresponding assessments were discussed indicator by indicator (Flint et al., 2016). The stakeholders agreed on most of the indicators and the respective assessments. If there were discrepancies, the final decision on the selection of indicators and their assessment was made by the scientists involved. For example, some stakeholders would have liked to see an assessment of the prevalence of lameness and joint injuries. However, the data quality for these 2 indicators is questionable. It is known from other studies that farmers significantly underestimate the prevalence of lameness in their dairy cows compared with professionally trained (external) individuals (Whay et al., 2002; Sarova et al., 2011). Therefore, it was decided to assess the corresponding management indicators (control of joint injuries and lameness). Nevertheless, recording and documenting animal-related indicators leads to a better awareness of the situation on one's own farm. This is more motivating for implementation of improved husbandry conditions and management factors to eliminate the causes of the recorded problems (Main et al., 2012). Because the DST is also an awareness tool, the prevalence of lameness and joint injuries was therefore recorded but not assessed. Feedback from farmers on the survey confirmed the approach: "Just filling out the questionnaire makes you think."

Table 2. Animal welfare indicators from Table 1 included in the Animal Welfare Index with weightings (in parentheses) and other methodological aspects

Total number of selected indicators	Number of merged indicators	Selected indicators (weighting of 0.25, 0.5, 0.33, 0.5, or 1)	Type of indicator[1]	Other methodological aspects		
				Data type[2]	Threshold[3]	Maximum achievable score
1	1	Freedom of movement for the dairy cows (husbandry system) (1)	b	d	g	2 or 1[4]
2–4	2	Access to outdoor climate stimuli for dairy cows (0.33); dry cows (0.33); young cattle (0.33)	b	d	g	2
5–6	3	Type of cubicle and cubicle floor covering (0.5); management of lying areas (0.5)	b, c	d	g	2
7	4	Ratio: resting areas to dairy cows (1)	b	e	g	2
8	5	Ratio: feeding areas to dairy cows (1)	b	e	g	2 or 1[4]
9–10	6	Number of drinking troughs (0.5); inspection and cleaning of drinking troughs (0.5)	b, c	e, d	g	2
11–12	7	Existence of cow brushes (0.5); facilities for thermoregulation (0.5)	b	d	g	2
13–15	8	Existence of calving pens (0.25); management of calving pens (0.25); existence of areas for sick dairy cows (0.5)	b, c	d, e	g	2
16	9	Proportion of difficult calvings	a	e	g	2
17–18	10	Proportion of dairy cows with ≤100,000 SCC/mL at average for the last 12 mo (0.5); way and frequency of maintenance of the milking system (0.5)	a, c	d, e	g	2
19	11	Strategy for the use of antibiotics for mastitis (1)	c	d	g	2
20	12	Control of joint injuries (1)	c	d	g	2
21	13	Performing cow-individual lameness control (1)	c	d	g	2
22	14	Frequency of prophylactic hoof care (1)	c	d	g	2
23–25	15	Metabolism profile of dairy cows: frequency of analysis (0.5); share of dairy cows with a fat:protein ratio ≥1.5 (0.25) or <1 (0.25) for the last year	c, a	d, e	g	2
26	16	Dairy cow mortality rate for the last year (1)	a	e	g	2
27	17	Herd care by external experts (1)	c	d	g	2
28	18	Care of resting areas for newborn calves (1)	c	d	g	2
29–30	19	Calf rearing management: supply of colostrum (0.5); supply of water (0.5)	c	d	f	2
31	20	Disbudding method (1)	c	d	f	2
32	21	Calf mortality rate for the last year (1)	a	e	g	2

[1] a = Animal based, b = resource based, c = management based.

[2] d = Qualitative data, e = quantitative data.

[3] f = Based on legal standards, g = based on expert.

[4] Farms with tiestalls can achieve a maximum of 1 point in this indicator.

In step 4, the questionnaire for measuring sustainability including animal welfare at the dairy farm level, which was revised from the previous versions, was field tested with dairy farmers. During the field test, on-farm interviews were conducted to check the feasibility and clarity of the questionnaire. Final adjustments to the questionnaire were made after the field test. Data have now been collected continuously since 2017.

These 4 steps were performed in earlier projects and in close consultation with other scientists specializing in animal welfare assessments and in several workshops with farmers and dairies (Lassen et al., 2014, 2015; Flint et al., 2016). In total, 49 experts have participated in the development of the animal welfare indicators and their assessments: 10 experts representing farmers, 7 dairy representatives, 2 processing industry representatives, 8 representatives of farmers' associations, 4 representatives of dairy associations, 1 representative of food retail associations, 2 representatives of animal welfare nongovernmental organizations, 5 representatives of agricultural extension services, and 10 scientists specializing in farm animal welfare.

As a result, we had a catalog of 46 animal welfare indicators available for analyses. These indicators have different meanings in terms of their ability to measure animal welfare. For example, indicators are included where it was clear that retailers require information about the indicator but that are viewed differently from a scientific point of view. Therefore, we conducted an expert survey to find out which of the DST animal welfare indicators should be used for a multidimensional assessment of animal welfare; that is, which indicators should be included in our AWI.

Aggregation of Animal Welfare Indicators into an Index

In 2019, we interviewed 11 scientists specializing in dairy cattle, including 2 veterinarians, to further condense the DST animal welfare indicators into our combined AWI. In total, we contacted 15 experts for the survey. Our experts were selected from all regions of Germany to cover the diverse structures of German dairy production. In a manner similar to a Delphi study, the expert survey was performed in written, anonymous form to avoid the influence of "super experts" (see Bélanger et al., 2012). For orientation, the experts were given a brief overview of the DST project. This means that the experts were aware of the purpose for which the indicators were collected; that is, what influenced the selection of the indicators (see section "Definition and Measurement of Animal Welfare"). The individual animal welfare indicators presented to the experts as

well as their assessment schemes (see Supplemental File S1, https://www.openagrar.de/receive/openagrar _mods_00081639). The expert survey was therefore dependent on how we assessed the respective indicators. We asked for an assessment of the importance of the indicators we selected as follows: "How would you weight the respective indicator in relation to animal welfare? Please assign values of 1 (low importance) to 10 (high importance) points."

We included all indicators with a mean importance of 8 points or more in our AWI with the exception of 3 indicators. The indicator "Access to pasture for dairy cows" (mean 7.3; see Supplemental File S1) was included because essentially a downward outlier caused the (lower) mean value of 7.3. Furthermore, due to the consistently positive effects on animal behavior, this indicator is often used in other assessment tools to evaluate animal welfare; therefore, in our opinion, it should not be omitted from the AWI. Moreover, instead of the prevalence of joint injuries (mean 8.9) and lameness (mean 9.6), we included the corresponding management indicators control of joint injuries (mean 7.3) and lameness (mean 8.2) in the AWI. We made this decision because, as mentioned above, data quality is questionable for these 2 indicators. The animal welfare indicators comprising our AWI are shown in Tables 1 and 2. In total, we included 32 indicators in the AWI (Table 2). Our AWI included indicators covering the Five Freedoms and the 4 categories of the Welfare Quality protocol (see Table 1), both of which serve as the basis for a holistic animal welfare assessment. Also, because of the large number of farms in this study, we believe this use of the AWI is appropriate for our purpose.

To aggregate the individual indicators into an index, we converted the information contained in the indicators into a standard, dimensionless scale. The assessment of the class characteristics of the individual animal welfare indicator J was expanded with a point scale P. A normalized value of 0 represents "level 0," a value of 2 "level 2," and a value of 3 "level 3." We decided to merge the "level 3" and "level 2" categories because the assessment scheme is applied specifically to the different indicators and is therefore used in different ways, so there is not always a "level 3." For example, compared with life in the wild, the behavior of dairy cows in any husbandry system is restricted, regardless of the structural or technical features of the husbandry system and management. With this in mind, there is no "level 3" assessment category for indicator 1 "Freedom of movement for the dairy cows (husbandry system)." If 3 points were assigned to the "level 3" category, some indicators (see Supplemental File S1) would have been weighted more heavily in the AWI from the outset,

whereas others would not. There are arguments in favor of merging these 2 assessment categories as well as arguments against it.

Scoring methodologies of tools for accessing sustainability or animal welfare often apply a "weight-and-sum" aggregation of indicators (de Olde et al., 2016). We followed essentially the same approach. The experts in our project noted that different indicators describe the same issue to some extent. However, this is not entirely avoidable if predominantly resource- and management-related indicators are used instead of animal-related indicators. This should be considered in the index formation. As a first step, we therefore combined several indicators that address similar animal welfare issues (see Table 2; indicators 2, 3, 6, 7, 8, 10, 15, and 19). The 32 selected indicators were thus merged into 21 indicators. Indeed, some indicators are included in the AWI with weights of 0.25, 0.33, or 0.5 (see Table 2). The combined 21 indicators are then included in the AWI with a weight of 1. Spoolder et al. (2003) argue that the selection and weighting of individual indicators, even if they are based on expert opinion and knowledge, ultimately always remain subjective. With that in mind, we think that a strength of our AWI is that it remains fairly simple and transparent.

Equation [1] shows how an individual farm's animal welfare score is computed according to the selected animal welfare indicators (Table 2):

$$AWI_i = \sum_{J=1}^{21} P_{j,i},\qquad [1]$$

for $i = 1, \ldots, N$ dairy farms, and $j = 1, \ldots, N$ animal welfare indicators, individual animal welfare indicator J, point scale P.

It can be assumed that farms with a higher AWI probably provide better conditions for good animal welfare than farms with a lower AWI if the management- and resource-based animal welfare indicators reported by farmers in the questionnaire are properly implemented. The minimum achievable score is 0 points, and the maximum achievable score is 42. The AWI can be used to evaluate farms with tiestalls and those with loose housing systems. With the exception of indicator 1 (restricted freedom of movement) and 8 (because there is no free choice of feeding place), farms with tiestalls can also achieve the maximum score for the respective indicators (Table 2).

Hypotheses and Analytical Framework

From the literature, a total of 5 core hypotheses were derived regarding the relationship between animal
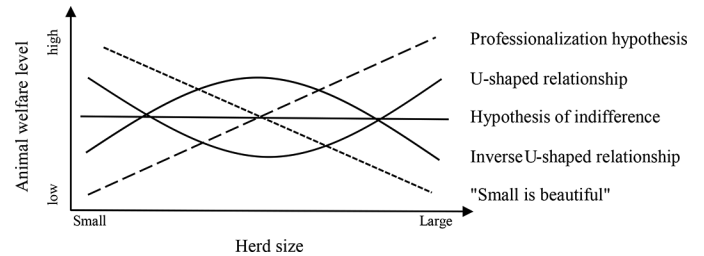

Figure 1. Possible links between animal welfare and herd size. Our illustration is based on Spiller et al. (2015).

welfare and herd size (Figure 1). But as will be shown below, there are theoretical counter-hypotheses to each of the 5 core hypotheses.

(1) **Professionalization hypothesis:** Growth and specialization of farms is accompanied by economies of scale and advantages in competence because, for example, larger farms seem to be more receptive to science-based recommendations (e.g., Hoe and Ruegg, 2006). Furthermore, larger farms are more likely to use advisory services specializing in animal health (e.g., Russell and Bewley, 2013), conduct routine veterinary herd health visits, or use monitoring systems (Beggs et al., 2015). Larger farms are also more likely to require and benefit from standard operating procedures and staff training to improve consistency and minimize human error (Hyde et al., 2011). According to these arguments, larger farms have a better animal welfare status. As a counter-argument, it is pointed out that although large farms are professionalized at the management level, in the actual animal care, they often rely on semi-skilled workers, who frequently exhibit competence deficits (Spiller et al., 2015). In addition, larger farms tend to have fewer workers per animal and are therefore at a disadvantage when it comes to providing individualized animal care (Robbins et al., 2016). Other authors point out that opponents of larger farms often assume a strict economic rationality of such farms with a consistent focus on cost reduction, even if animal welfare deteriorates as a result (Hess et al., 2014).

(2) **"Small is beautiful":** According to this hypothesis, a multifunctional orientation, high motivation, and greater competence of the farm manager and family members working on the farm can improve the animal welfare situation on smaller family farms without employed labor. In contrast, it is argued that small farms with several branches of business frequently have

knowledge deficits and little scope for innovation due to business management problems (Spiller et al., 2015). Furthermore, several studies have shown that farmers on smaller farms tend to be more stressed (Simkin et al., 1998), which, according to Robbins et al. (2016), may also put animal welfare at risk.

(3) **U-shaped relationship:** Small farms may have advantages in animal welfare (see the "small is beautiful" hypothesis) that are lost with increasing farm size. Medium-sized family farms in particular (without outside labor) can reach the limits of their working capacity, which could pose a risk to animal welfare. From a certain farm size onward, however, animal welfare may increase again, as increasing professionalization of farm processes can be ensured (see "professionalization hypothesis"; Spiller et al., 2015). Ultimately, this hypothesis is, to some extent, a combination of the "small is beautiful" and "professionalization" hypotheses.

(4) **Inverse U-shaped relationship:** Although small farms may have disadvantages due to low specialization and lack of training, large farms often do not succeed in recruiting appropriately trained or motivated staff, so that medium-sized enterprises can ensure the best animal welfare (Spiller et al., 2015).

(5) **Indifference hypothesis:** Overall, the links between animal welfare and farm size are not very pronounced and are camouflaged by management's competence and size-independent issues of farm structure, so that, on average, no reliable relationship can be established (e.g., Robbins et al., 2016).

In the overall view of all 5 core hypotheses, it becomes apparent that no definitive main hypothesis can be derived. Thus, the relationship between animal welfare and herd size at the dairy farm level is examined empirically in the present study using all 5 hypotheses. Although this study focuses on the influence of herd size, it is ultimately not the only factor influencing the animal welfare level on individual farms. Recent scientific studies have shown that the role of farmers, for example, is central to improving animal welfare. Therefore, the factors that influence farmers' thinking may ultimately affect individual animal welfare on the farm. In particular, these include farmers' knowledge and the cost implications of farm animal welfare (Balzani and Hanlon, 2020). However, personality (e.g., Adler et al., 2019), values (e.g., Hansson and Lagerkvist, 2016; Hansson et al., 2018; Heise and Theuvsen, 2018), com-

munication with their veterinarian and farm advisors (e.g., Becker et al., 2014; Väärikkälä et al., 2018), as well as time and management (Balzani and Hanlon, 2020) influence perceptions of animal welfare. We therefore define animal welfare level AWI of an individual farm i as a function (f) of $\mathbf{S}$ (herd size), $\mathbf{F}$ (farm characteristics), $\mathbf{P}$ (personal incentives), $\mathbf{M}$ (good farm management), and $\mathbf{E}$ (economic influence):

$$AWI_i = f(\mathbf{S}_i,\ \mathbf{F}_i,\ \mathbf{P}_i,\ \mathbf{M}_i,\ \mathbf{E}_i). \qquad [2]$$

The following variables (in italics) were identified from the available data set, which can be interpreted as approximations of the respective vectors of explanatory variables.

Farm Characteristics (F). It is assumed that *full-time farms* as well as *specialized dairy farms* have a higher AWI due to the lower opportunity costs that farms face for time spent with the cattle (Hess et al., 2014). For example, the scarce production factor time can be used for the cows and not elsewhere, for example, in arable farming. Furthermore, we assume that *organic farms* have a higher AWI due to their voluntary adherence to higher standards (e.g., Spiller et al., 2015). Against the background of agglomeration effects and spillover effects, we assume that *farms in regions with a high proportion of grassland* (= dairy intensive region) have a higher AWI, because, for example, knowledge exchange between farms is easier to achieve due to spatial proximity (Lindena and Hess, 2018). We also assume that *farms that are determined to continue dairy production* have a higher AWI, because these farms are preparing for the future and are therefore addressing the increasingly important issue of animal welfare. Hansen and Østerås (2019), for example, found that determination to continue production was associated with better animal welfare indirectly through farm expansion.

Personal Incentives (P). Skills, knowledge, and motivation of stockpeople (e.g., dairy farmers) to effectively care for and manage their animals, as well as farmers' attitudes and behavior, are integral to animal welfare (e.g., Waiblinger et al., 2002; Hemsworth, 2018). We assume, therefore, that the *farmers' education, use of advisory services specialized in dairy farming and animal health*, or *participation in off-farm training* results in a higher AWI. Furthermore, we assume that *farmers who are more satisfied with their personal working situation* on the farm also make efforts with animal welfare (e.g., Hansen and Østerås, 2019) because these farmers are less stressed. We also assume that the *farmers' age* (experience through age; e.g., Owusu-Sekyere et al., 2022, or young and motivated) may be relevant for the

willingness to adopt animal welfare measures and thus influence the AWI level.

Good Farm Management (M). The more time and energy farms can dedicate to the animals and the less stress the farmers have (e.g., Hansen and Østerås, 2019), the higher the AWI. The *number of dairy cows per labor unit, total workload on the farm,* and *utilizing robot milking* were therefore included in the estimation.

Economic Influence (E). Both positive ("animal welfare pays off"; e.g., Telldahl et al., 2019) and negative ("animal welfare costs"; i.e., necessary investment costs may prevent dairy farmers from changing their production system toward more animal welfare; Lagerkvist et al., 2011) effects of *economic satisfaction* on the AWI can be assumed.

Methods

To analyze the link between animal welfare and herd size, we used ordinary least squares (**OLS**) regression. After adding a stochastic error term (u) to the model in Eq. [2], the following estimation equation was obtained:

$$AWI_i = \beta_0 + \beta'_S lnS_i + \beta'_F F_i + \beta'_P P_i + \beta'_M M_i + \beta'_E E_i + u_i,$$

[3a]

where β represents the estimated coefficients.

The model in Eq. [3a] tests the hypotheses "Professionalization," "Indifference," and "Small is beautiful" (Figure 1). The logarithmic transformation of *herd size* (lnS_i) is useful to limit the negative effects of outlier values in the data. Furthermore, the estimated coefficient can be interpreted as the effect of a 1% change in *herd size* on AWI. The "U-shaped" and "Inverse U-shaped" hypotheses can be approximated by removing the log and adding polynomial terms for the explanatory variable S:

$$AWI_i = \beta_0 + \beta'_S S_i + \beta'_S S_i^2 + \beta'_S S_i^3 + \beta'_F F_i + \beta'_P P_i + \beta'_M M_i + \beta'_E E_i + u_i.$$

[3b]

Selection and weighting of individual indicators are crucial for the outcome of the animal welfare assessment and thus also for our regression results. To assess the sensitivity of our results, we therefore re-estimated the model in Eq. [3a] with differently computed indices as the dependent variable. In addition to the AWI formation shown in Eq. [1], we formulated our index alternatively as follows:

(1) using a 4-point scale (level 3 = 3 points, level 2 = 2 points, level 1 = 1 point, and level 0 = 0 points) instead of our implemented 3-point scale where level 2 and level 3 were combined

(2) indicators were grouped according to the 4 dimensions of the Welfare Quality protocol (Table 1), with each dimension contributing 25% to the index

(3) indicators were grouped according to the Five Freedoms (Table 1), with each dimension contributing 20% to the index

Using this approach, we attempted to check whether and how the results of the regression change when our index is calculated differently; that is, with different weights for the individual indicators.

We also conducted a quantile regression (Koenker and Hallock, 2001). Quantile regression models the relationship between animal welfare and the farm's characteristics (lnS_i, F_i, P_i, M_i, E_i) using the conditional quantile, which allows evaluation of the specific effects of these characteristics on different groups of farms clustered on their level of AWI. The hypothesis to be tested was that coefficients of herd size on animal welfare vary according to quantiles (0.05, 0.25, 0.5, 0.75, 0.95).

RESULTS AND DISCUSSION

Descriptive Statistics of Covariates

The descriptive statistics of the explanatory variables are summarized in Table 3. Even though our sample is a convenience sample (see section "Data from the German Dairy Sustainability Tool"), it closely approximates the diverse structures of dairy farming in Germany: our sample consisted mainly of conventional (98%; 2% were organic) full-time farms (97%) specialized in dairy production (93%; see Table 3). The average dairy farmer in our sample had 141 ha of agricultural land, of which 73% was grassland (permanent, temporary, or both). The average dairy farmer was 47 years old. In our sample, 12% held a university degree, which is in line with the German farming population, of which 11% have a university degree (BMEL, 2021). Agricultural college degrees ("Fachschule"), on the other hand, were significantly overrepresented in our sample (36% in our sample and 14% in the German farming population; BMEL, 2021). Seventeen percent of our farm managers indicated that they would very probably give up milk production in the next 10 years, which is in line with the observed structural change in Germany (Forstner and Nieberg, 2019). Mean milk yield was 8,810 kg per dairy cow per year (in Germany, 8,250 kg per dairy cow per year; BMEL, 2021). In total, 55% of the farms in our sample allowed cows to graze, accounting for 39% of lactating dairy cows. For comparison, in Germany

Table 3. Descriptive statistics of the explanatory variables (n = 3,085)

Explanatory variable	Mean	% of farms	SD	Minimum	Maximum
Herd size (no. of dairy cows per farm)	122		118	7	2,900
Specialized dairy farm (D[1])	0.93		0.25	0	1
Farming represents the main activity (D)	0.97		0.18	0	1
Conventional farming (D)	0.98		0.14	0	1
Share of grassland of total utilized agricultural area					
a) <30%	—	11.09	—	—	—
b) 30% to <70%	—	54.65	—	—	—
c) ≥70%	—	34.26	—	—	—
Expected existence of dairy production in 10 yr?					
a) "Yes, definitely"	—	34.07	—	—	—
b) "Rather likely"	—	34.42	—	—	—
c) "I cannot estimate"	—	14.72	—	—	—
d) "Rather unlikely"	—	10.76	—	—	—
e) "Certainly not"	—	6.03	—	—	—
Age of farmer (yr)	47		10	19	83
Agricultural education					
a) Vocational training with a "master agriculture" degree or agricultural college or university degree	—	82.53	—	—	—
b) Vocational training (agriculture)	—	11.93	—	—	—
c) No agricultural education	—	5.54	—	—	—
Use of advisory services specialized in dairy, last 3 yr (D)	0.52		0.50	0	1
Use of advisory services specialized in animal health, last 3 yr (D)	0.44		0.50	0	1
Farmer participating in off-farm training, last 3 yr (D)	0.67		0.47	0	1
Farmer's satisfaction with the personal work situation					
a) Very satisfied	—	3.01	—	—	—
b) Satisfied	—	34.59	—	—	—
c) Rather satisfied	—	31.86	—	—	—
d) Rather dissatisfied	—	21.33	—	—	—
e) Dissatisfied	—	7.13	—	—	—
f) Very dissatisfied	—	2.07	—	—	—
Number of dairy cows per labor unit	42		20	3.5	280
Workload total farm …					
a) is manageable	—	9.85	—	—	—
b) is mostly very high	—	61.00	—	—	—
c) is consistently too high	—	29.14	—	—	—
Farms having robot milking (D)	0.18		0.38	0	1
Modernization investments in dairy farm, last 5 yr (D)	0.86		0.35	0	1
Economic satisfaction, last 3 yr					
a) Very satisfied	—	1.53	—	—	—
b) Satisfied	—	16.69	—	—	—
c) Rather satisfied	—	24.67	—	—	—
d) Rather dissatisfied	—	31.83	—	—	—
e) Dissatisfied	—	15.88	—	—	—
f) Very dissatisfied	—	9.40	—	—	—

[1]D = dummy variable, which has a value of 1 if an attribute is present and 0 otherwise.

overall, 45% of dairy farms allow cows to graze, accounting for 42% of all dairy cows (BMEL, 2021).

Table 4 shows that all herd size classes, according to the official herd size statistics of Germany, were included. The smallest dairy farm in the data set had 7 dairy cows, and the largest had 2,900 dairy cows (Table 3). Our sample thus covered a wide range of dairy farm sizes. Mean herd size in our sample was larger (122 dairy cows per farm) than the average German herd size (mean 70 dairy cows per farm; Table 4). In particular, the following 3 aspects influenced the higher average in our study: farms in our sample were overrepresented in the northwestern region and underrepresented in the southern and central regions (Table 4). Due to different dairy farm structures in the northwestern and southern

and central regions, medium-sized farms were thus overrepresented and smaller farms underrepresented (Table 4). Furthermore, many participating dairies reported that it was very difficult to motivate smaller dairy farms to complete the questionnaire. This was observed in all regions; in all regions, our herd size was larger than the German average (Table 4). In addition, we observed that smaller farms, in particular, tended to fill in the questionnaire incompletely; as a result, the mean farm size in our sample was significantly higher (122 dairy cows per farm) than the original data set (93 dairy cows per farm). In our analyses, we only included farms (n = 3,085) that had provided all relevant information for the regression analysis. A total of about 130 individual variables had to be assessed. Even if only one piece of

Table 4. Comparison of the population of German dairy farms and our sample

| Dairy cows per farm[1] | Proportion of farms (%) | | Region | Proportion of farms (%) | | Mean herd size | |
	Germany[2] (n = 54,787)	Our sample[3] (n = 3,085)		Germany[2] (n = 54,787)	Our sample[3] (n = 3,085)	Germany[2] (n = 54,787)	Our sample[3] (n = 3,085)
1–9	15.5	0.2	Germany			70	122
10–19	11.1	0.9	Northwestern[4]	30.3	66.0	93	122
20–49	27.3	10.6	Southern and central[5]	63.7	30.8	47	94
50–99	26.8	40.0	Eastern[6]	6.0	3.2	198	388
100–199	14.0	36.1					
200–499	4.3	10.8					
≥500	1.0	1.4					

[1]Herd size classes according to the Federal Statistical Office.
[2]Data from the Federal Statistical Office (Destatis, 2021).
[3]Data from the Dairy Sustainability Tool.
[4]Federal States: Schleswig Holstein, Lower Saxony, Hamburg, Bremen, North Rhine-Westphalia.
[5]Federal States: Hesse, Rheinland-Palatinate, Baden-Wuerttemberg, Bavaria, Saarland.
[6]Federal States: Berlin, Brandenburg, Mecklenburg-Western Pomerania, Saxony, Saxony-Anhalt, Thuringia.

information was missing, this resulted in exclusion from our analysis. This affected approximately 4,000 farms from the original data set (n = 7,297). However, the fact that smaller farms participate less in surveys was also reported by Petersen and Hess (2018), who conducted a Germany-wide survey with dairy farmers during the same period but on a topic unrelated to animal welfare. We therefore assumed that the reluctance of smaller farms to respond was unrelated to animal welfare issues on their farm, but rather was due to the opportunity cost of time. Against the background of the continuous structural change toward fewer but larger farms in Germany, the bias toward larger farms in our sample is not necessarily a disadvantage (see also Schreiner and Hess, 2017).

Animal Welfare Situation on Farms

The farms in our data set achieved individual animal welfare scores of between 15.8 and 41.5 points; the mean AWI was 30.6. This indicates that, on average, 73% of the maximum possible AWI value of 42 points was reached. The animal welfare situation, as judged by our AWI, was at an intermediate level, with potential for improvement. In comparison, Gieseke et al. (2018), who analyzed a data set of 80 German dairy farms using the Welfare Quality protocol, classified the majority of farms as having "enhanced" (30%) or "acceptable" (66%) animal welfare.

Considering the construct validity of the AWI, we might suspect a circularity problem because herd size as an explanatory variable could be closely related or even coincide with the indicators used in our AWI (dependent variable). The variable "Freedom of movement for the dairy cows" used in the AWI provides a good illustration of this: tiestalls, with their more restricted animal movement, compared with loose housing systems, coincide with small farms (Figure 2). It should not be surprising, therefore, that smaller farm sizes are associated with a lower AWI. Conversely, compared with farms with loose housing (most of the larger farms), farms with tiestalls usually score well on indicators such as the cow-to-resting area ratio, as each cow has its own cubicle. In this context, it should not be surprising if, conversely, larger dairy farms achieve a lower AWI. Overall, we assume that there is no circularity problem.

The AWI was plotted against the number of dairy cows per farm, also distinguishing farms by region and housing system (Figure 2). On average, farms with loose housing achieved a higher AWI (mean 30.9) than farms with tiestalls (mean 27.1). However, the scattering of AWI values showed that both high and low AWI values can be achieved, regardless of the husbandry system, because certain indicators such as resting areas per cow are independent of the dairy barn system and can substitute for other factors. Furthermore, it can be seen that differences in AWI within regions are greater than those between regions (Figure 2). Mean values for AWI were 30.5 in the northwestern region, 30.7 in the southern region, and 32.1 in the eastern region. With respect to AWI, variation was not only evident within regions but also within herds of the same size. This result was also observed, for example, by Gieseke et al. (2018), who found large differences in welfare levels between dairy farms with similar herd sizes, despite using a different set of indicators (the Welfare Quality protocol).

At first glance, it is difficult to derive a link between animal welfare and herd size from these graphs. However, even though the number of observations for the larger farms was small, none of the 6 plots contained
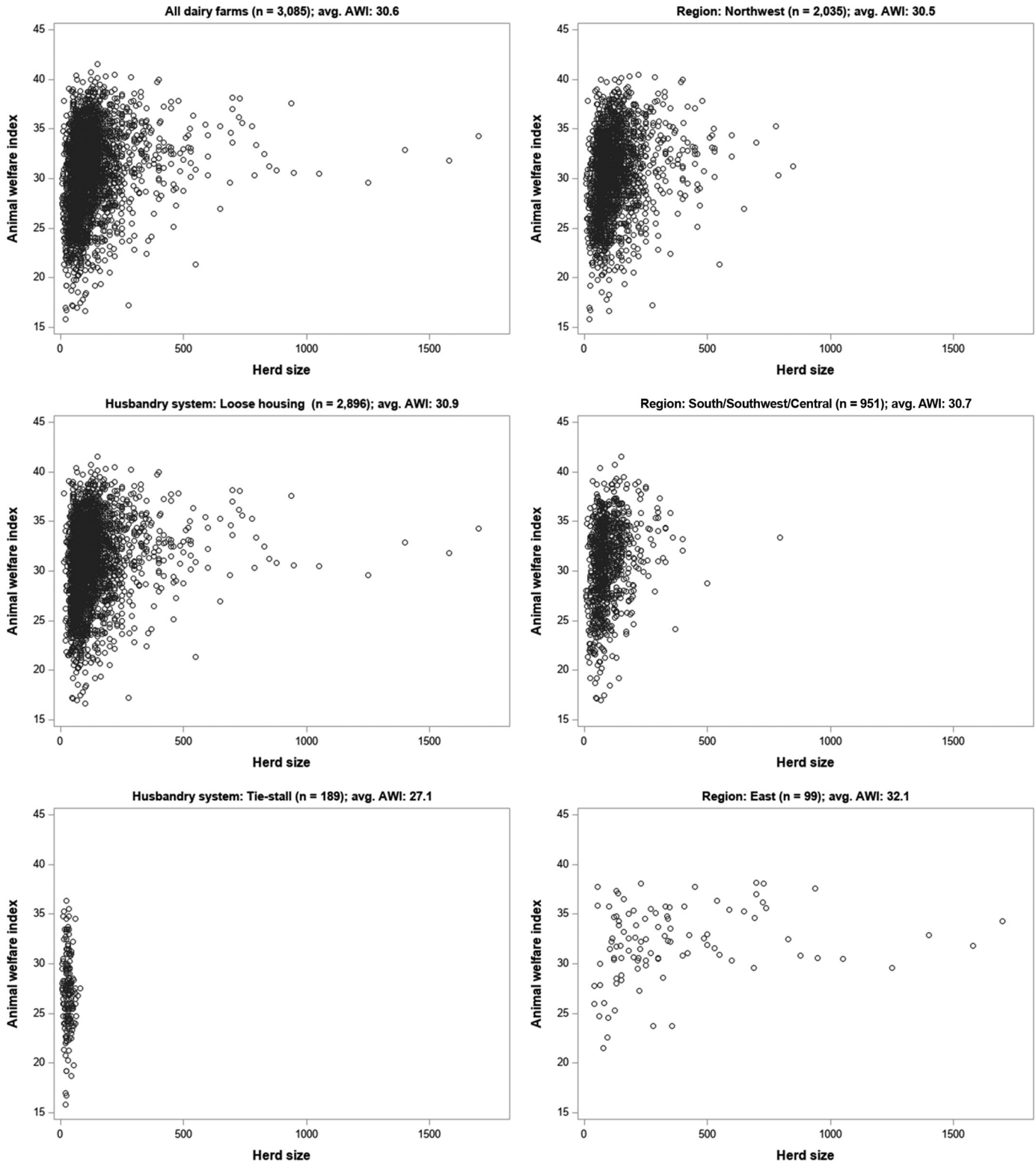
Figure 2. Scatterplots of animal welfare index (AWI) versus herd size of German dairy farms by housing system and region. One dairy farm with 2,900 dairy cows (AWI = 39) is not shown, for scaling reasons.

any observations in the lower right-hand corner (i.e., high number of animals and low AWI). Thus, there was initially no empirical indication of the largest farms in our sample having the lowest AWI.

Econometric Specifications

Equation [3a] was initially estimated as a regression model with OLS estimator. Residuals had a normal distribution, and robust standard errors were applied and were clustered by German federal states to reflect different variances of the residuals that are due to regional differences in farm types. We refrained from clustering standard errors into smaller administrative units because the federal states already reflect the regional differences quite well: a comparison of the means and standard deviations of the explanatory variables shows that they are scattered among the federal states. Furthermore, the values of the variance inflation factors showed that multicollinearity was not an issue. For 2 very similar dummies of *farmers' satisfaction with the personal work situation* (satisfied, rather satisfied), values of 13 and 14 were obtained, but all other values were clearly below the usually applied threshold of 10. In addition, the Ramsey specification test showed that the model was appropriately specified ($P = 0.039$; tested polynomials: squares and cubes).

To analyze the "U-shaped" and "inverse U-shaped" hypotheses, we nevertheless (even though the Ramsey test rejected it) added polynomial terms for the explanatory variable S (see Eq. [3b]). The estimated coefficient for farm size (S) was found to be positive. In addition, a negative (positive) coefficient was found for S^2 (S^3), which implies an exponentially increasing curve for farm size overall. This means, that our AWI increased with increasing farm size.

In addition to the OLS regression Eq. [3a], we conducted a quantile regression to model the specific effects of a farm's characteristics ($\ln S_i$, F_i, P_i, M_i, E_i) on different groups of farms clustered on their level of AWI. The estimates for farm size for the different quantiles (0.05, 0.25, 0.5, 0.75, 0.95) barely differed from the OLS estimates presented in Table 5. This implies that coefficients of farm size on animal welfare do not vary according to AWI quantiles.

To obtain an indication of the presence of potential endogeneity bias in the estimation, the variables *farmers' satisfaction with the personal work situation*, *economic satisfaction*, *participation in off-farm training*, and *use of dairy or animal health specialized advisory services* were alternately excluded from the OLS estimation as described in Eq. [3a]. Auer and Rottmann (2012) conclude from strongly deviating coefficients of the OLS estimates (with and without the respective potentially endogenous variable) that endogeneity can influence the estimation results. Because the coefficients changed only minimally, we concluded that endogeneity was not a problem in our model. We also estimated Eq. [3a] without potentially influential observations, without outliers concerning farm size and based on more homogeneous subsamples (e.g., without organic farms or farms with tiestalls). However, qualitative findings were similar for all alternative model types; therefore, we focused on the interpretation of regression as described in Eq. [3a].

Regression Results

Our results indicated that larger farms tended to achieve a higher AWI than smaller farms (Table 5). The upper and lower limits of the "compatibility interval" (Amrhein et al., 2019), showing the interval estimates that are most compatible with our data and our model, were strictly positive. A dairy farm that is 1% larger than the average dairy farm in our sample scored c.p. (all else being equal) 0.0069 AWI points more on average. In another example, a farm with 134 dairy cows (i.e., 10% more dairy cows than the sample average) achieved c.p. 0.066 AWI points more in our model. A farm that was twice as large (+100%) as the average farm in our sample achieved c.p. 0.693 AWI points more in our model; thus, achieving a total AWI of 31.3 points. In the context of our AWI scoring system, this means that a farm with 244 dairy cows is not even 1 point better in one indicator than our average farm with 122 dairy cows, which achieves 30.6 AWI points. The effect size is therefore very small.

Our results suggest that, with respect to the relationship between animal welfare and herd size, the "Professionalization hypothesis" (Figure 1) is valid; that is, a positive relationship exists between farm size and animal welfare for dairy farms, which Wellbrock et al. (2009) and Hess et al. (2014) also found, even though their data were for pig farms. Hess et al. (2014) attribute their results to specialization and lower opportunity costs of larger farms. The results of Gieseke et al. (2018) do not indicate a linear relationship between herd size and animal welfare (as an overall score) in German dairy farms. However, their results are not completely different from ours: among the individual principles (each principle is composed of several indicators), a positive relationship was found for the principle "good feeding," which the authors attribute to management practices on larger dairy farms. The other principles (good housing, good health, and appropriate behavior) were not affected by herd size (Gieseke et al., 2018).

Table 5. Results of the ordinary least squares (OLS) estimations to explain the animal welfare index (n = 3,085)[1]

Explanatory variable	Coefficient	95% CI or compatibility interval	P-value
Herd size (log)	0.693	(0.322, 1.063)	0.000
Specialized dairy farm (D[2])	1.138	(0.580, 1.696)	<0.000
Farming represents the main activity (D)	−0.865	(−1.850, 0.120)	0.085
Conventional farming (D)	−0.663	(−1.173, −0.153)	0.011
Share of grassland of total utilized agricultural area			
a) 30% to <70%	0.307	(−0.278, 0.893)	0.304
b) ≥70%	0.243	(−0.640, 1.125)	0.590
Expected existence of dairy production in 10 years?			
a) "Yes, definitely"	0.546	(0.012, 1.079)	0.045
b) "Rather likely"	0.455	(−0.163, 1.074)	0.149
c) "I cannot estimate"	0.131	(−0.456, 0.718)	0.663
d) "Rather unlikely"	−0.211	(−0.925, 0.503)	0.562
Age of farmer (years)	0.006	(−0.008, 0.019)	0.404
Agricultural education			
a) Vocational training with a "master agriculture" degree or agricultural college or university degree	0.619	(0.185, 1.054)	0.005
b) Vocational training (agriculture)	−0.116	(−0.494, 0.261)	0.545
Use of advisory services, dairy, last 3 yr (D)	1.010	(0.770, 1.251)	<0.000
Use of advisory services, animal health, last 3 yr (D)	1.302	(1.185, 1.420)	<0.000
Farmer participating in off-farm training, last 3 yr (D)	1.336	(0.997, 1.674)	<0.000
Farmer's satisfaction with the personal work situation			
a) Very satisfied	0.533	(−0.101, 1.167)	0.100
b) Satisfied	0.153	(−0.526, 0.832)	0.659
c) Rather satisfied	−0.044	(−0.726, 0.637)	0.898
d) Rather dissatisfied	0.013	(−0.556, 0.583)	0.963
e) Dissatisfied	−0.150	(−0.987, 0.687)	0.725
Number of dairy cows per labor unit	−0.012	(−0.017, −0.007)	<0.000
Workload total farm . . .			
a) is frequently very high	−0.015	(−0.372, 0.343)	0.936
b) is consistently too high	−0.053	(−0.441, 0.336)	0.790
Farms having robot milking (D)	0.914	(0.587, 1.241)	<0.000
Modernization investments in dairy farm, last 5 yr (D)	0.665	(0.439, 0.891)	<0.000
Economic satisfaction, last 3 yr			
a) Very satisfied	0.438	(−0.296, 1.172)	0.242
b) Satisfied	−0.214	(−0.833, 0.405)	0.498
c) Rather satisfied	−0.302	(−0.811, 0.208)	0.246
d) Rather dissatisfied	0.027	(−0.421, 0.474)	0.907
e) Dissatisfied	−0.136	(−0.415, 0.143)	0.340
Const	24.384	(22.950, 25.819)	<0.000

[1]$R^2 = 0.218$; adjusted $R^2 = 0.210$.

[2]D = dummy variable, which has a value of 1 if an attribute is present and 0 otherwise.

As a robustness check, we calculated Eq. [3a] again with different index variants: (1) using a 4-point scale instead of a 3-point scale, (2) weighted according to the 4 dimensions of the Welfare Quality protocol, and (3) weighted according to the Five Freedoms. We found no fundamental differences in the results for variants (1) and (2). In the 2 index variants, the coefficient for herd size remained positive: for variant (1), the coefficient was 0.742 (95% confidence interval of 0.256 to 1.228 and a P-value of 0.003, Table 6 and Appendix 1 Table A2); for variant (2), we found an (even) smaller effect size: the coefficient for herd size was 0.231 (confidence interval of 0.054 to 0.406 and a P-value of 0.011, see Table 6 and Appendix 2 Table A3). Variant (3) provided a different picture: the coefficient for herd size was −0.054, with a confidence interval of −0.165 to 0.057 and a P-value of 0.343 (Table 6 and Appendix

2 Table A4). This result is not surprising. Considering the assignment of our indicators to the Five Freedoms (see Table 1), we see that the categories "freedom from hunger and thirst" and "freedom from fear and distress" are each served by only few indicators. Among others, these 2 freedoms include the indicators "Ratio: resting areas to dairy cows," "Ratio: feeding areas to dairy cows," and "Number of drinking troughs and inspection and cleaning of drinking troughs." In particular, these indicators were generally well met by farms with ties-talls (as each cow has usually its own cubicle) and these farms are usually relatively small in terms of herd size (see Figure 2). Because all Five Freedoms are equally weighted in the index (20% each), the above-mentioned indicators related to space per cow receive more weight in the index. This explains why the estimated coefficient on herd size is negative (statistically insignificant)

Table 6. Robustness check: Results of the ordinary least squares (OLS) estimations of herd size to explain the animal welfare index (n = 3,085) using alternative approaches to calculate the index

Index formation	Coefficient	95% CI or compatibility interval	P-value
Using a 3-point scale	0.693	(0.322, 1.063)	0.000
Using a 4-point scale	0.742	(0.256, 1.228)	0.003
Indicators grouped according to the 4 Welfare Quality dimensions	0.231	(0.054, 0.406)	0.011
Indicators grouped according to the Five Freedoms	−0.054	(−0.165, 0.057)	0.343

under this particular aggregation scheme of the animal welfare indicators.

Because our sample represents larger farms, the question arises of whether the relation between animal welfare and herd size would be different if our sample included a larger number of smaller farms. A closer look at the descriptive statistics shows that smaller dairy farms were more likely to be found in the southern regions and that southern German dairy farms are underrepresented in our data set thus far (Table 4). Figure 2 shows that dairy farms in southern Germany achieved AWI values in the same range as the AWI in eastern and northwestern Germany. Against this background, we do not expect that our results would change profoundly if the analysis included a greater number of smaller dairy farms.

Based on our findings, larger dairy herds are not necessarily associated with poorer animal welfare. Therefore, enforced restrictions on herd size do not lead to improved animal welfare. However, despite the positive relationship between animal welfare and herd size in our study, there is potential for better animal welfare on small, medium-sized, and large dairy farms, as we found heterogeneity in all herd size classes in our data set with respect to AWI. Robbins et al. (2016) point to another interesting aspect of the "small is beautiful" debate: "the over-simplified view that animal welfare is better on smaller farms may create complacency among small farmers (allowing welfare problems to persist), and fails to focus efforts on specific welfare challenges that need to be resolved on farms of all sizes." (Robbins et al., 2016, page 21).

Other variables (Table 5) also imply an effect on the dependent variable. In brief, farmers having a higher *agricultural education degree*, *using advisory services*, or *participating in off-farm training* tend to have a higher AWI. Furthermore, several results suggest that farms on which more time is spent with dairy cows (i.e., have fewer *dairy cows per labor unit*) or have lower opportunity costs to spend time with their dairy cows (e.g., *specialized dairy farms* or *farms using robotic milking systems*) tend to achieve a higher AWI. In addition,

future-oriented farms (*modernization investments* or *dairy production anticipated for more than 10 years to come*) tend to have a higher AWI. As with *herd size*, the effect size of all independent variables was small. For example, if a dairy farmer has made a modernization investment in the last 5 years, the AWI was c.p. 0.66 points higher than that of farmers who have not made an investment.

Overall, only about one-fifth of the AWI (adjusted R^2 = 0.210) could be explained by *herd size* and other vectors as described in Eq. [3a]. Four-fifths of the animal welfare described by our AWI must therefore be due to other influencing factors that we do not capture or capture with insufficient detail with our model, such as the farmer's inner attitude toward animal welfare or the farmer's ability to handle his or her cows.

CONCLUSIONS

The main objective of this study was to investigate the relationship between animal welfare and herd size of German dairy farms using an AWI, which was developed as part of a larger project on sustainability in dairy farming. Theoretically, positive, negative, and curvilinear relationships between animal welfare and herd size are conceivable, but also a statistically nonsignificant effect for the relationship between these 2 variables might occur. Our regression results showed that larger farms tended to achieve a higher AWI that smaller farms, but the effect size was very small. Nevertheless, contrary to the assumption in public discourse, larger dairy herds were not necessarily associated with poorer animal welfare. In all herd size classes, we found a large variation of AWI between herds and thus potential for better animal welfare on small, medium-sized, and large dairy farms. Our results strengthen the evidence that herd size has little, if any, effect on farm-specific animal welfare levels. Therefore, when animal welfare is discussed in public and in politics, the emphasis should be on implementing animal welfare measures on farms, with less focus on herd size or politically enforced herd size restrictions.

ACKNOWLEDGMENTS

REFERENCES

Adler, F., R. Christley, and A. Campe. 2019. Invited review: Examining farmers' personalities and attitudes as possible risk factors for dairy cattle health, welfare, productivity, and farm management: A systematic scoping review. J. Dairy Sci. 102:3805–3824. https://doi.org/10.3168/jds.2018-15037.

Amrhein, V., D. Trafimow, and S. Greenland. 2019. Inferential statistics as descriptive statistics: There is no replication crisis if we don't expect replication. Am. Stat. 73(Suppl. 1):262–270. https://doi.org/10.1080/00031305.2018.1543137.

Andrade, S. B., and I. Anneberg. 2014. Farmers under pressure. Analysis of the social conditions 485 of cases of animal neglect. J. Agric. Environ. Ethics 27:103–126. https://doi.org/10.1007/s10806-013-9456-9.

Auer, B., and H. Rottmann. 2012. Statistik und Ökonometrie für Wirtschaftswissenschaftler: Eine anwendungsorientierte Einführung. 2. Aufl. Gabler Verlag.

Balzani, A., and A. Hanlon. 2020. Factors that influence farmers' views on farm animal welfare: A semi-systematic review and thematic analysis. Animals (Basel) 10:1524. https://doi.org/10.3390/ani10091524.

Becker, J., M. Reist, and A. Steiner. 2014. Factors influencing the attitudes of cattle veterinarians, farmers, and claw trimmers towards the pain associated with the treatment of sole ulcers and the sensitivity to pain of dairy cows. Vet. J. 200:38–43. https://doi.org/10.1016/j.tvjl.2014.01.016.

Beggs, D. S., A. D. Fisher, E. C. Jongman, and P. E. Hemsworth. 2015. A survey of Australian dairy farmers to investigate animal welfare risks associated with increasing scale of production. J. Dairy Sci. 98:5330–5338. https://doi.org/10.3168/jds.2014-9239.

Beggs, D. S., E. C. Jongman, P. H. Hemsworth, and A. D. Fisher. 2019. The effects of herd size on the welfare of dairy cows in a pasture-based system using animal- and resource-based indicators. J. Dairy Sci. 102:3406–3420. https://doi.org/10.3168/jds.2018-14850.

Bélanger, V., A. Vanasse, D. Parent, G. Allard, and D. Pellerin. 2012. Development of agri-environmental indicators to assess dairy farm sustainability in Quebec, Eastern Canada. Ecol. Indic. 23:421–430. https://doi.org/10.1016/j.ecolind.2012.04.027.

BMEL (Bundesministerium für Ernährung und Landwirtschaft). 2021. Statistisches Jahrbuch über Ernährung, Landwirtschaft und Forsten: Jahrbuch 2020. Bundesinformationszentrum Landwirtschaft (BZL).

Botreau, R., M. Bonde, A. Butterworth, P. Perny, M. B. M. Bracke, J. Capdeville, and I. Veissier. 2007. Aggregation of measures to produce an overall assessment of animal welfare. Part 1: A review of existing methods. Animal 1:1179–1187. https://doi.org/10.1017/S1751731107000535.

Busch, G., M. Gauly, and A. Spiller. 2018. Opinion paper: What needs to be changed for successful future livestock farming in Europe? Animal 12:1999–2001. https://doi.org/10.1017/S1751731118001258.

Czekaj, T. G., A. S. Nielsen, A. Henningsen, B. Forkman, and M. Lund. 2013. The Relationship Between Animal Welfare and Economic Outcome at the Farm Level. IFRO Report 222, Department of Food and Resource Economics, University of Copenhagen.

de Olde, E. M., F. W. Oudshoorn, C. A. Sørensen, E. A. Bokkers, and I. J. de Boer. 2016. Assessing sustainability at farm-level: Lessons learned from a comparison of tools in practice. Ecol. Indic. 66:391–404. https://doi.org/10.1016/j.ecolind.2016.01.047.

Destatis. 2021. Land- und Forstwirtschaft, Fischerei: Viehbestand - 3. November 2020. Fachserie 3 4.1. Accessed Jan. 25, 2021. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Landwirtschaft-Forstwirtschaft-Fischerei/Tiere-Tierische-Erzeugung/Publikationen/Downloads-Tiere-und-tierische-Erzeugung/viehbestand-2030410215325.xlsx?__blob=publicationFile.

Deutscher Tierschutzbund e.V. 2021. Richtlinien Milchkühe 2021: Kriterienkatalog für die Haltung und Behandlung von Milchkühen. Accessed June 29, 2021. https://www.tierschutzlabel.info/richtlinien.

European Commission. 2013. Structure and dynamics of EU farms: Changes, trends and policy relevance. EU Agricultural Economics Brief No. 9, Directorate-General for Agriculture and Rural Development, European Commission, 2013.

European Commission. 2016. Attitudes of Europeans towards animal welfare: November-December 2015. Special Eurobarometer 442. Accessed Jan. 05, 2019. https://data.europa.eu/data/datasets/s2096_84_4_442_eng?locale=de.

FAWC (Farm Animal Welfare Council). 1979. Press Statement. News release. December 5, 1979. Accessed Sep. 12, 2022. https://webarchive.nationalarchives.gov.uk/ukgwa/20110704164609mp_/http://www.fawc.org.uk/pdf/fivefreedoms1979.pdf.

Flint, L., H. Kuhnert, B. Laggner, B. Lassen, H. Nieberg, and R. Strohm. 2016. Prozess Nachhaltige Milcherzeugung - Entwicklung eines Nachhaltigkeitsmoduls zur Erfassung und Bewertung von Nachhaltigkeitskriterien auf Milchviehhaltenden Betrieben. Thünen Working Paper 54. Thünen-Institute. https://doi.org/10.3220/WP1459335513000.

Forstner, B., and H. Nieberg. 2019. Agrarstrukturentwicklung - Bestimmende Faktoren und Perspektiven. Landentwicklung Aktuell 24:65–67.

Fraser, D. 2008. Understanding animal welfare. Acta Vet. Scand. 50(Suppl. 1):S1–7. https://doi.org/10.1186/1751-0147-50-S1-S1.

Gieseke, D., C. Lambertz, and M. Gauly. 2018. Relationship between herd size and measures of animal welfare on dairy cattle farms with freestall housing in Germany. J. Dairy Sci. 101:7397–7411. https://doi.org/10.3168/jds.2017-14232.

Hansen, B. G., and O. Østerås. 2019. Farmer welfare and animal welfare—Exploring the relationship between farmer's occupational well-being and stress, farm expansion and animal welfare. Prev. Vet. Med. 170:104741–8. https://doi.org/10.1016/j.prevetmed.2019.104741.

Hansson, H., and C. J. Lagerkvist. 2016. Dairy farmers' use and nonuse values in animal welfare: Determining the empirical content and structure with anchored best-worst scaling. J. Dairy Sci. 99:579–592. https://doi.org/10.3168/jds.2015-9755.

Hansson, H., C. J. Lagerkvist, and K. M. Vesala. 2018. Impact of personal values and personality on motivational factors for farmers to work with farm animal welfare: A case of Swedish dairy farmers. Anim. Welf. 27:133–145. https://doi.org/10.7120/09627286.27.2.133.

Heise, H., and L. Theuvsen. 2018. German dairy farmers' attitudes toward farm animal welfare and their willingness to participate in animal welfare programs: A cluster analysis. Int. Food Agribus.

Manag. Rev. 21:1121–1136. https://doi.org/10.22434/IFAMR2017.0066.

Hemme, T. 2020. IFCN Dairy Report 2020. International Farm Comparison Network.

Hemsworth, P. H. 2018. Key determinants of pig welfare: Implications of animal management and housing design on livestock welfare. Anim. Prod. Sci. 58:1375–1386. https://doi.org/10.1071/AN17897.

Hess, S., L. A. Bolos, R. Hoffmann, and Y. Surry. 2014. Is animal welfare better on small farms? Evidence from veterinary inspections on Swedish farms. Paper presented at the 14th International Congress of the European Association of Agricultural Economists (EAAE) in Ljubljana, Slovenia, Aug. 26–29, 2014. Accessed Sep. 12, 2022. https://ageconsearch.umn.edu/record/182781/.

Hoe, F. G. H., and P. L. Ruegg. 2006. Opinions and practices of Wisconsin dairy producers about biosecurity and animal well-being. J. Dairy Sci. 89:2297–2308. https://doi.org/10.3168/jds.S0022-0302(06)72301-3.

Hyde, J., S. A. Cornelisse, and L. A. Holden. 2011. Human resource management on dairy farms: Does investing in people matter? Econ. Bull. 31:208–217.

Koenker, R. K. F. H., and K. F. Hallock. 2001. Quantile regression. J. Econ. Perspect. 15:143–156. https://doi.org/10.1257/jep.15.4.143.

Lagerkvist, C. J., H. Hansson, S. Hess, and R. Hoffman. 2011. Provision of farm animal welfare: Integrating productivity and non-use values. Appl. Econ. Perspect. Policy 33:484–509. https://doi.org/10.1093/aepp/ppr037.

Lassen, B., H. Nieberg, H. Kuhnert, and J. Sanders. 2014. Status-quo-Analyse ausgewählter Nachhaltigkeitsaspekte der Milcherzeugung in Niedersachsen. Thünen Working Paper 28. Thünen-Institute. https://doi.org/10.3220/WP_28_2014.

Lassen, B., H. Nieberg, H. Kuhnert, J. Sanders, and R. Schleenbecker. 2015. Status-quo-Analyse ausgewählter Nachhaltigkeitsaspekte der Milcherzeugung in Schleswig-Holstein. Thünen Working Paper 43. Thünen-Institute. https://doi.org/10.3220/WP1439978006000.

Lindena, T., and S. Hess. 2018. Besteht ein Zusammenhang zwischen nachhaltiger Milcherzeugung und wirtschaftlicher Zufriedenheit der Betriebsleiter? Ergebnisse aus Schleswig-Holstein. In Agrar- und Ernährungswirtschaft zwischen Ressourceneffizienz und gesellschaftlichen Erwartungen, edited by Johannes Sauer. 1. Auflage, 29–40. Schriften der Gesellschaft für Wirtschafts- und Sozialwissenschaften des Landbaues e.V 53. Landwirtschaftsverlag.

Main, D. C. J., K. A. Leach, Z. E. Barker, A. K. Sedgwick, C. M. Maggs, N. J. Bell, and H. R. Whay. 2012. Evaluating an intervention to reduce lameness in dairy cattle. J. Dairy Sci. 95:2946–2954. https://doi.org/10.3168/jds.2011-4678.

Marchant-Forde, J. N. 2015. The science of animal behavior and welfare: Challenges, opportunities, and global perspective. Front. Vet. Sci. 2:16. https://doi.org/10.3389/fvets.2015.00016.

Mason, G. J., and M. Mendl. 1993. Why is there no simple way of measuring animal welfare? Anim. Welf. 2:301–319.

Mazurek, M., D. J. Prendiville, M. A. Crowe, I. Veissier, and B. Earley. 2010. An on-farm investigation of beef suckler herds using an animal welfare index (AWI). BMC Vet. Res. 6:55. https://doi.org/10.1186/1746-6148-6-55.

Otten, N. D., L. R. Nielsen, P. T. Thomsen, and H. Houe. 2014. Register-based predictors of violations of animal welfare legislation in dairy herds. Animal 8:1963–1970. https://doi.org/10.1017/S1751731114001918.

Owusu-Sekyere, E., H. Hansson, and E. Telezhenko. 2022. Use and non-use values to explain farmers' motivation for the provision of animal welfare. Eur. Rev. Agric. Econ. 49:499–525. https://doi.org/10.1093/erae/jbab012.

Petersen, J., and S. Hess. 2018. The future of milk delivery relations from the perspective of German farmers. Ger. J. Agric. Econ. 67:80–94.

Pfeiffer, J., A. Gabriel, and M. Gandorfer. 2021. Understanding the public attitudinal acceptance of digital farming technologies: A nationwide survey in Germany. Agric. Human Values 38:107–128. https://doi.org/10.1007/s10460-020-10145-2.

Robbins, J. A., M. A. G. von Keyserlingk, D. Fraser, and D. M. Weary. 2016. Invited Review: Farm size and animal welfare. J. Anim. Sci. 94:5439–5455. https://doi.org/10.2527/jas.2016-0805.

Roesch, A., G. Gaillard, J. Isenring, C. Jurt, N. Keil, and T. Nemecek. 2016. Umfassende Beurteilung der Nachhaltigkeit von Landwirtschaftsbetrieben. Agroscope science Umwelt 33. Agroscope.

Russell, R. A., and J. M. Bewley. 2013. Characterization of Kentucky dairy producer decision-making behavior. J. Dairy Sci. 96:4751–4758. https://doi.org/10.3168/jds.2012-6538.

Sarova, R., I. Stehulova, P. Kratinova, P. Firla, and M. Spinka. 2011. Farm managers underestimate lameness prevalence in Czech dairy herds. Anim. Welf. 20:201–204.

Sauer, J., and U. Latacz-Lohmann. 2015. Investment, technical change and efficiency: Empirical evidence from German dairy production. Eur. Rev. Agric. Econ. 42:151–175. https://doi.org/10.1093/erae/jbu015.

Schreiner, J. A., and S. Hess. 2017. The role of non-use values in dairy farmers' willingness to accept a farm animal welfare programme. J. Agric. Econ. 68:553–578. https://doi.org/10.1111/1477-9552.12203.

Scott, E. M., A. M. Nolan, and J. L. Fitzpatrick. 2001. Conceptual and methodological issues related to welfare assessment: A framework for measurement. Acta Agric. Scand. A Anim. Sci. 51(Suppl. 30):5–10. https://doi.org/10.1080/090647001316922983.

Simkin, S., K. Hawton, J. Fagg, and A. Malmberg. 1998. Stress in farmers: A survey of farmers in England and Wales. Occup. Environ. Med. 55:729–734. https://doi.org/10.1136/oem.55.11.729.

Sørensen, J. T., and D. Fraser. 2010. On-farm welfare assessment for regulatory purposes: Issues and possible solutions. Livest. Sci. 131:1–7. https://doi.org/10.1016/j.livsci.2010.02.025.

Spiller, A., M. Gauly, A. Balmann, J. Bauhus, R. Birner, W. Bokelmann, O. Christen, S. Entenmann, H. Grethe, U. Knierim, U. Latacz-Lohmann, J. Martinez, H. Nieberg, M. Qaim, F. Taube, B.-A. Tenhagen, and P. Weingarten. 2015. Wege zu einer gesellschaftlich akzeptierten Nutztierhaltung. Berichte über Landwirtschaft (Sonderheft Nr. 221). https://doi.org/10.12767/buel.v0i221.82.

Spoolder, H., G. de Rosa, B. Hörning, S. Waiblinger, and F. Wemelsfelder. 2003. Integrating parameters to assess on-farm welfare. Anim. Welf. 12:529–534.

Stott, A. W., B. Vosough Ahmadi, C. M. Dwyer, B. Kupiec, C. Morgan-Davies, C. E. Milne, S. Ringrose, P. Goddard, K. Phillips, and A. Waterhouse. 2012. Interactions between profit and welfare on extensive sheep farms. Anim. Welf. 21:57–64. https://doi.org/10.7120/096272812X13345905673683.

Telldahl, C., H. Hansson, and U. Emanuelson. 2019. Modelling animal health as a production factor in dairy production—A case of low somatic cell counts in Swedish dairy agriculture. Livest. Sci. 230:103840. https://doi.org/10.1016/j.livsci.2019.103840.

Väärikkälä, S., S. M. Artukka, L. Hänninen, and M. Nevas. 2018. Finnish cattle and pig farmers' perceptions of animal welfare inspections. Anim. Welf. 27:369–377. https://doi.org/10.7120/09627286.27.4.369.

Vanhonacker, F., W. Verbeke, E. van Poucke, and F. A. Tuyttens. 2008. Do citizens and farmers interpret the concept of farm animal welfare differently? Livest. Sci. 116:126–136. https://doi.org/10.1016/j.livsci.2007.09.017.

Waiblinger, S., U. Knierim, and C. Winckler. 2001. The development of an epidemiologically based on-farm welfare assessment system for use with dairy cows. Acta Agric. Scand. A Anim. Sci. 51(Suppl. 30):73–77. https://doi.org/10.1080/090647001316923108.

Waiblinger, S., C. Menke, and G. Coleman. 2002. The relationship between attitudes, personal characteristics and behaviour of stockpeople and subsequent behaviour and production of dairy cows. Appl. Anim. Behav. Sci. 79:195–219. https://doi.org/10.1016/S0168-1591(02)00155-7.

Welfare Quality Consortium. 2009. Welfare Quality Assessment protocol for cattle. Welfare Quality Consortium, Lylystad, Netherlands. Accessed Jan. 07, 2019. http://www.welfarequalitynetwork.net/media/1088/cattle_protocol_without_veal_calves.pdf.

Wellbrock, W., S. J. Oosting, B. B. Bock, B. Antunović, and G. Kralik. 2009. Harmonization of welfare standards for the protection of pigs with the EU rules: The case of Croatia. Ital. J. Anim. Sci. 8(Suppl. 3):21–38.

Whay, H. R., D. Main, L. E. Green, and A. Webster. 2002. Farmer perception of lameness prevalence. Pages 355–358 in Proc. 12th Int. Symp. Lameness in Ruminants, Orlando, FL. J. K. Shearer, ed.

Zimmermann, A., and T. Heckelei. 2012. Structural change of European dairy farms—A cross-regional analysis. J. Agric. Econ. 63:576–603. https://doi.org/10.1111/j.1477-9552.2012.00355.x.

ORCIDS

T. Lindena ● https://orcid.org/0000-0003-0723-1541
S. Hess ● https://orcid.org/0000-0002-4597-7654

APPENDIX 1

Table A1. Extract from the questionnaire: Existence of areas for sick dairy cows

Do you have a separate area in the farm for sick cows or for cows with rehabilitation needs, or an area that can be set up at short notice if necessary?	☐ Yes, a rehabilitation area is available ☐ Yes, a rehabilitation area can be set up at short notice ☐ No
If so, how many areas are available or can be set up at short notice?	☐ One rehabilitation area ☐ More than one rehabilitation area
How much space does a sick cow have on average in the rehabilitation area?	☐ Less than 10 m^2 ☐ 10 to <15 m^2 ☐ 15 m^2 or more
For how many cows is a separate rehabilitation area available or quickly set up at short notice?	For maximum _____ cows

Table A2. Results of the ordinary least squares (OLS) estimations to explain the animal welfare index (n = 3,085) – index formation using a 4-point scale[1]

Explanatory variable	Coefficient	95% CI or compatibility interval	P-value
Herd size (log)	0.742	(0.256, 1.228)	0.003
Specialized dairy farm (D[2])	1.261	(0.623, 1.899)	<0.000
Farming represents the main activity (D)	−1.389	(−2.635, −0.143)	0.029
Conventional farming (D)	−0.792	(−1.510, −0.073)	0.031
Share of grassland of total utilized agricultural area			
a) 30% to <70%	0.160	(−0.531, 0.850)	0.652
b) ≥70%	−0.036	(−1.172. 1.101)	0.951
Expected existence of dairy production in 10 years?			
a) "Yes, definitely"	0.506	(−0.289, 1.301)	0.212
b) "Rather likely"	0.316	(−0.523, 1.154)	0.461
c) "I cannot estimate"	−0.039	(−0.834, 0.756)	0.923
d) "Rather unlikely"	−0.505	(−1.383, 0.373)	0.260
Age of farmer (yr)	0.004	(−0.010, 0.018)	0.568
Agricultural education			
a) Vocational training with a "master agriculture" degree or agricultural college or university degree	0.596	(0.013, 1.179)	0.045
b) Vocational training (agriculture)	−0.086	(−0.599, 0.427)	0.743
Use of advisory services, dairy, last 3 yr (D)	1.297	(0.968, 1.626)	<0.000
Use of advisory services, animal health, last 3 yr (D)	1.701	(1.525, 1.877)	<0.000
Farmer participating in off-farm training, last 3 yr (D)	1.681	(1.263, 2.099)	<0.000
Farmer's satisfaction with the personal work situation			
a) Very satisfied	1.230	(0.212, 2.248)	0.018
b) Satisfied	0.638	(−0.354, 1.629)	0.207
c) Rather satisfied	0.383	(−0.636, 1.403)	0.461
d) Rather dissatisfied	0.372	(−0.507, 1.251)	0.407
e) Dissatisfied	0.234	(−0.956, 1.423)	0.700
Number of dairy cows per labor unit	−0.015	(−0.024, −0.006)	0.001
Workload total farm …			
a) is frequently very high	−0.015	(−0.471, 0.441)	0.949
b) is consistently too high	−0.069	(−0.361, 0.506)	0.756
Farms having robot milking (D)	1.345	(0.906, 1.784)	<0.000
Modernization investments in dairy farm, last 5 yr (D)	0.888	(0.615, 1.161)	<0.000
Economic satisfaction, last 3 yr			
a) Very satisfied	0.206	(−0.324, 0.735)	0.446
b) Satisfied	−0.403	(−1.329, 0.523)	0.394
c) Rather satisfied	−0.536	(−1.116, 0.044)	0.070
d) Rather dissatisfied	−0.212	(−0.763, 0.339)	0.451
e) Dissatisfied	−0.340	(−0.696, 0.017)	0.062
Constant	28.038	(26.341, 29.735)	<0.000

[1]4-point scale (level 3 = 3 points, level 2 = 2 points, level 1 = 1 point, and level 0 = 0 points) instead of our implemented 3-point scale where level 2 and level 3 were combined. $R^2 = 0.208$; adjusted $R^2 = 0.201$.

[2]D = dummy variable, which has a value of 1 if an attribute is present and 0 otherwise.

Table A3. Results of the ordinary least squares (OLS) estimations to explain the animal welfare index (n = 3,085) – index formation using indicators grouped according to the 4 dimensions of the Welfare Quality protocol (good health, good housing, good feeding, appropriate behavior), with each dimension contributing 25% to the index[1]

Explanatory variable	Coefficient	95% CI or compatibility interval	P-value
Herd size (log)	0.231	(0.054, 0.406)	0.011
Specialized dairy farm (D[2])	0.615	(0.247, 0.983)	0.001
Farming represents the main activity (D)	−0.548	(−1.078, −0.017)	0.043
Conventional farming (D)	−0.458	(−0.874, −0.043)	0.031
Share of grassland of total utilized agricultural area			
a) 30% to <70%	0.143	(−0.220, 0.505)	0.441
b) ≥70%	0.102	(−0.430, 0.633)	0.707
Expected existence of dairy production in 10 years?			
a) "Yes, definitely"	0.295	(−0.060, 0.651)	0.103
b) "Rather likely"	0.231	(−0.162, 0.625)	0.250
c) "I cannot estimate"	0.118	(−0.281, 0.517)	0.561
d) "Rather unlikely"	−0.118	(−0.541, 0.305)	0.581
Age of farmer (years)	0.003	(−0.005, 0.012)	0.445
Agricultural education			
a) Vocational training with a "master agriculture" degree or agricultural college or university degree	0.265	(−0.048, 0.578)	0.097
b) Vocational training (agriculture)	−0.032	(−0.301, 0.227)	0.816
Use of advisory services, dairy, last 3 yr (D)	0.529	(0.423, 0.636)	<0.000
Use of advisory services, animal health, last 3 yr (D)	0.808	(0.740, 0.876)	<0.000
Farmer participating in off-farm training, last 3 yr (D)	0.775	(0.586, 0.965)	<0.000
Farmer's satisfaction with the personal work situation			
a) Very satisfied	0.431	(−0.108, 0.971)	0.117
b) Satisfied	0.220	(−0.283, 0.724)	0.391
c) Rather satisfied	0.060	(−0.462, 0.583)	0.821
d) Rather dissatisfied	0.119	(−0.279, 0.516)	0.558
e) Dissatisfied	0.047	(−0.507, 0.602)	0.867
Number of dairy cows per labor unit	−0.006	(−0.010, −0.003)	<0.000
Workload total farm …			
a) is frequently very high	−0.012	(−0.238, 0.214)	0.918
b) is consistently too high	−0.053	(−0.305, 0.200)	0.681
Farms having robot milking (D)	0.551	(0.378, 0.723)	<0.000
Modernization investments in dairy farm, last 5 yr (D)	0.359	(0.209, 0.509)	<0.000
Economic satisfaction, last 3 yr			
a) Very satisfied	0.059	(−0.360, 0.478)	0.783
b) Satisfied	−0.359	(−0.712, −0.005)	0.047
c) Rather satisfied	−0.318	(−0.607, −0.029)	0.031
d) Rather dissatisfied	−0.085	(−0.280, 0.110)	0.391
e) Dissatisfied	−0.126	(−0.276, 0.025)	0.102
Const	15.177	(14.173, 16.182)	<0.000

[1]$R^2 = 0.182$; adjusted $R^2 = 0.174$

[2]D = dummy variable, which has a value of 1 if an attribute is present and 0 otherwise.

Table A4. Results of the ordinary least squares (OLS) estimations to explain the animal welfare index (n = 3,085) – index formation: indicators grouped according to the Five Freedoms, with each dimension contributing 20%

Explanatory variable	Coefficient	95% CI or compatibility interval	P-value
Herd size (log)	−0.054	(−0.165, 0.057)	0.343
Specialized dairy farm (D[2])	0.381	(0.131, 0.630)	0.003
Farming represents the main activity (D)	−0.237	(−0.529, 0.055)	0.112
Conventional farming (D)	−0.511	(−0.707, −0.315)	<0.000
Share of grassland of total utilized agricultural area			
a) 30% to <70%	0.134	(−0.106, 0.374)	0.275
b) ≥70%	0.108	(−0.227, 0.443)	0.527
Expected existence of dairy production in 10 years?			
a) "Yes, definitely"	0.203	(0.032, 0.437)	0.090
b) "Rather likely"	0.183	(−0.052, 0.417)	0.126
c) "I cannot estimate"	0.080	(−0.111, 0.271)	0.413
d) "Rather unlikely"	−0.054	(−0.282, 0.174)	0.643
Age of farmer (years)	−0.001	(−0.006, 0.005)	0.822
Agricultural education			
a) Vocational training with a "master agriculture" degree or agricultural college or university degree	0.200	(0.050, 0.348)	0.009
b) Vocational training (agriculture)	−0.007	(−0.098, 0.085)	0.888
Use of advisory services, dairy, last 3 yr (D)	0.300	(0.236, 0.364)	<0.000
Use of advisory services, animal health, last 3 yr (D)	0.427	(0.371, 0.483)	<0.000
Farmer participating in off-farm training, last 3 yr (D)	0.474	(0.335, 0.613)	<0.000
Farmer's satisfaction with the personal work situation			
a) Very satisfied	0.274	(−0.027, 0.576)	0.075
b) Satisfied	0.118	(−0.260, 0.496)	0.540
c) Rather satisfied	0.059	(−0.331, 0.450)	0.766
d) Rather dissatisfied	0.103	(−0.244, 0.450)	0.561
e) Dissatisfied	0.008	(−0.392, 0.409)	0.967
Number of dairy cows per labor unit	−0.005	(−0.007, −0.003)	<0.000
Workload total farm …			
a) is frequently very high	−0.012	(−0.125, 0.102)	0.842
b) is consistently too high	−0.056	(−0.177, 0.065)	0.362
Farms having robot milking (D)	0.383	(0.214, 0.551)	<0.000
Modernization investments in dairy farm, last 5 yr (D)	0.274	(0.198, 0.351)	<0.000
Economic satisfaction, last 3 yr			
a) Very satisfied	0.109	(−0.145, 0.364)	0.400
b) Satisfied	−0.129	(−0.374, 0.116)	0.301
c) Rather satisfied	−0.175	(−0.384, 0.035)	0.103
d) Rather dissatisfied	−0.019	(−0.183, 0.145)	0.821
e) Dissatisfied	−0.068	(−0.220, 0.083)	0.377
Const	11.948	(11.296, 12.600)	<0.000

[1]$R^2 = 0.145$; adjusted $R^2 = 0.137$.

[2]D = dummy variable, which has a value of 1 if an attribute is present and 0 otherwise.