

Cost-efficiency analysis of animal welfare improvements in Dutch growing-finishing pigs



Author: Anne Hoogstra

April 2019



WAGENINGEN
UNIVERSITY & RESEARCH



100years
1918 — 2018

Cost-efficiency analysis of animal welfare improvements in Dutch growing-finishing pigs

MSc Thesis Business Economics

Author: Anne Hoogstra
Reg. number: 960805361100
University: Wageningen University
MSc programme: Animal Sciences
Chair group: Business Economics Group

Supervisors: dr.ir. H.W. Saatkamp (*Business Economics Group, Wageningen University*)
ir. H. M. Vermeer (*Wageningen Livestock Research, Wageningen Research*)

Examiner: prof.dr.ir. M. P. M. Meuwissen (*Business Economics Group, Wageningen University*)

Thesis code: BEC-80436
Period: September 2018 - March 2019
Date: 03-04-2019



Summary

During the past decades, public concern increased regarding animal welfare (AW) in many European countries, among which, the Netherlands. As a response, European countries gradually set better AW standards for all kinds of livestock. Besides that, the market reacted to the demand for animal friendlier products by creating the organic market and later, all kind of systems in between conventional and organic. Because of the public concern, farmers are increasingly expected to produce beyond the legal minimum standards in order to keep their license to produce. When it comes to the decision-making for a certain system type, the cost-efficiency of a system change is a valuable source of information for farmers and other stakeholders. This research has the objective to gain insight in the cost-efficiency of animal welfare improvements in the Dutch growing-finishing pig sector.

In order to reach this objective, a four-step approach was developed. This approach is derived from the method of (Gocsik *et al.*, 2016). The first step in the research is an inventory of the different pig production systems in the Netherlands and based on that a selection for analysis. This study focuses on the systems: conventional, Beter Leven 1-star small and big groups respectively, free-range and organic. The estimation of the AW level of the selected production systems was done using the Welfare Quality®(WQ) Assessment protocol for pigs (Welfare Quality®, 2009). For the reason that the scores for the WQ measures and the system attributes are not one to one comparable, a method had to be developed to link these two. These linkages were weighted to obtain allocation formulas. The production costs (variable and fixed costs) of the selected systems were calculated with the help of a deterministic model described in (Gocsik *et al.*, 2015) in step 3. Some of the variables used in the model were updated with literature and consulting experts. Step 4 contains the calculation of the cost-efficiency ($\Delta WQ / \Delta TC$) ratio for each system. This step includes a sensitivity analysis in order to test the strength of the conclusions and to check the influence of price fluctuations.

Results of this research show that WQ index scores of these systems were very similar. The difference between the lowest and highest score is only 6.0%. The organic system scores best with a total WQ score of 1348 after that in decreasing order the free-range system (1336), the 1-star system (1329) and finally the conventional system (1272). The three system attributes that showed the highest contribution were: outdoor access, stocking density and bedding. Different trends were observed between the systems for the measure scores and the system attribute scores. The variable costs showed an increasing trend from conventional to the organic system. Fixed costs are lowest in the free-range system, followed by the conventional system, Beter Leven 1 star big groups, Beter Leven 1 star small groups and finally the organic system. The variable and fixed costs added, resulted in the total costs to be approximately twice as high in the organic system (€295) compared to the other systems with total costs of €139 for conventional, €143 for 1-star small groups, €144 for 1-star big groups and €160 for free-range. Overall, the highest cost-efficiency ($\Delta WQ / \Delta Costs = 12.9$) can be observed converting from conventional to the Beter Leven 1-star system. All the other comparisons showed a considerably lower cost-efficiency: from conventional to free-range (3.0) and from conventional to organic (0.5). The sensitivity analysis showed that the ranking of costs was quite robust. Likewise, when correcting for the higher prices for feed and piglets in the organic system, the ranking did not change even though the differences were smaller. The welfare scores of 1-star, free-range and organic system overlapped and showed no clear ranking. This was expected following the results.

Concluding, it can be stated that the welfare level in all systems are very similar, while the costs did increase considerably converting to higher systems. Therefore, the highest cost-efficiency can be obtained converting from the conventional system to the 1-star system. The implications and recommendations for future research are discussed.

Table of Contents

Summary.....	i
List of tables	iii
List of figures	iii
List of abbreviations and definitions	iv
1 Introduction.....	1
2 Methodology	3
2.1 General approach	3
2.2 Step 1: Selection of suitable production systems	4
2.3 Step 2: Calculation of the welfare quality index score for production systems and individual measures.....	5
2.3.1 Welfare Quality® Assessment protocol for pigs	5
2.3.2 Decomposition of production systems into system attributes.....	9
2.3.3 Linkages and weights between system attributes and welfare measures	9
2.3.4 Data.....	16
2.3.5 Calculation scores.....	17
2.4 Step 3: Calculation of production costs.....	19
2.4.1 Economic model.....	19
2.4.2 Data.....	19
2.5 Step 4: Analyse the cost-efficiency of the pig production system.....	19
2.5.1 Calculation of cost-efficiency	19
2.5.2 Sensitivity analysis	19
3 Results	21
3.1 Welfare Quality index scores.....	21
3.2 Production costs	23
3.3 Cost-efficiency of production systems	29
3.4 Sensitivity analysis.....	30
4 Discussion	33
4.1 Results.....	33
4.2 Methodology	34
4.2.1 Linking and weighting system attributes to welfare measures	34
4.2.2 Data and sensitivity	35
4.2.3 Assessment of animal welfare using WQ	35
4.2.4 Conclusion	36
4.3 Implications	36
4.4 Future research and system development.....	37
5 Conclusion.....	38
Acknowledgements	39
References.....	40
Appendix I. Boxplots of selected welfare measure scores per system types.....	45
Appendix II. Formulas used in economic model from (Gocsik <i>et al.</i> , 2015).....	47

List of tables

Table 1: Overview of the main existing pig production systems in the Netherlands, ranked according to the Beter Leven quality mark system.	4
Table 2: Overview of the selected Dutch pig production system types and their characteristics regarding AW	5
Table 3: Display of the selected WQ measures and the corresponding criteria and principles	9
Table 4: Matrix showing the linkages between the selected Welfare Quality measures and the selected system attributes.....	10
Table 5: Matrix showing the assigned weights to the linkages between the WQ measures and the system attributes.....	11
Table 6: Technical inputs, costs and prices used in the economic model	20
Table 7: Mean welfare score per welfare measure and per system type.....	21
Table 8: Mean attributional welfare scores per system attribute and per system type	22
Table 9: Variable costs (€/delivered pig) per system type	24
Table 10: Overview of the proportion of feed- and piglet costs relative to the total variable costs.....	25
Table 11: Fixed costs (€/delivered pig) per system type	26
Table 12: Change in WQ Index score (Δ WQ) and production costs (Δ TC in €/delivered pig), their corresponding change in % and cost-efficiency (CE) per system type.....	29
Table 13: Results of the sensitivity analysis on WQ Index scores and production costs (€/delivered pig) per system type.....	30
Table 14: Affected variable costs when corrected for organic feed and piglet prices in € per delivered pig of the production systems	31

List of figures

Figure 1: Development of the number of integrally sustainable pig houses for the different certification systems on the 1st of January of the indicated year (Peet et al., 2018).	1
Figure 2: Overview of the research steps.....	3
Figure 3: The principles, criteria and measures of the WQ protocol for pigs.....	6
Figure 4: The approach of WQ for integrating the data on the different measures to an overall assessment of the animal unit.....	6
Figure 5: Overview of the WQ score calculations.....	17
Figure 6: Mean attributional welfare scores per system attribute and per system type.....	22
Figure 7: Change in attributional welfare score of the system types relative to the conventional system	23
Figure 8: Variable costs (€/delivered pig) per system type	24
Figure 9: Variable costs (€/delivered pig) excluding feed- and piglet costs.....	25
Figure 10: Fixed costs (€/delivered pig) per system type	26
Figure 11: Overview of fixed costs (€/delivered pig) per system type excluding labour costs	27
Figure 12: Total production costs (€/delivered pig) per system type	28
Figure 13: Plot of the total welfare quality score and the total production costs of the different system types and their corresponding cost-efficiency	29
Figure 14: Results of the sensitivity analysis on WQ Index scores and production costs (€/delivered pig) per system type including the WQ standard errors (vertical) and the best- and worst case costs scenario (horizontal).....	31
Figure 15: Plot of the total welfare quality score and the total production costs of the different system types plus the corrected organic system and their corresponding cost-efficiency.....	32
Figure 16: Maslow pyramid of needs	35

List of abbreviations and definitions

Attributional WQ score	Contribution of a specific attribute to the WQ Index score; the attributional WQ score is calculated by summing the multiplied welfare scores per measure and per barn with their defined weights
AW	Animal welfare; the well-being of animals
CE	Cost-efficiency: the difference in welfare score divided by the corresponding total production costs
QBA	Qualitative Behaviour Assessment; used to assess positive emotional state
System attribute	Specific aspect of a production system
System type/ production type	Type of pig production system; in the study, conventional, 1-star Beter Leven, free-range and organic were the included system types. A production system is composition of (AW influencing) attributes
TC	Total costs, consist out of variable and fixed costs of the production types
Welfare measure	A measure to assess a specific welfare criterion (Welfare Quality®, 2009)
Welfare Quality® Index score	The mean WQ score of a specific system type
WQ	Welfare Quality®

1 Introduction

Over 50 years ago, society in Europe focussed on abundance and cheapness of food after experiencing hunger during WWII. The demand for animal originated products grew, which resulted in the development of highly-efficient animal production systems (Lassen, Sandøe and Forkman, 2006). These highly efficient systems had a negative side effect, namely, reduced animal welfare (AW). Consequently, public concern increased regarding AW in many European countries, among which, the Netherlands (Appleby, 2004). As a response, European countries gradually set better AW standards for all kinds of livestock. Besides that, the market reacted to the demand for animal friendlier products by creating the organic market. The last decades, all kind of production systems in between conventional and organic were created to offer the costumers a wider range of animal friendlier products in different prices categories (Michelsen *et al.* 1999; Sørensen *et al.* 2006). Some countries, like the Netherlands, already set higher minimum legal standards for AW than the European standards (Veissier *et al.*, 2008).

Currently, in the Dutch meat sector, one can distinguish the following market segments regarding AW: conventional, that produces according to the minimal legal Dutch AW standards, top-segment with the highest AW standards, that includes organic and free-range systems, and a middle segment with the systems in between (Gocsik *et al.*, 2014). One way for consumers to see how animal-friendly a product is, is with the Beter Leven quality mark developed by the animal protection. This quality mark uses a star system, with a range from zero (conventional) to three stars (organic) and is visible with a sticker on the packaging (Beter Leven Keurmerk, 2019a). Figure 1 shows the use of different quality marks in the pig sector in the Netherlands over the past years, among which the Beter Leven quality mark.

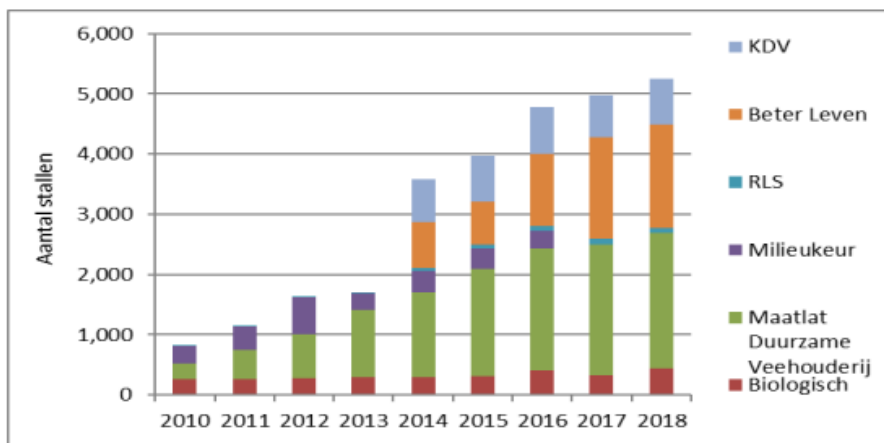


Figure 1: Development of the number of integrally sustainable pig houses for the different certification systems on the 1st of January of the indicated year (Peet *et al.*, 2018).

A scientifically-based way to assess the animal welfare in animal production systems is with the help of the welfare quality assessment protocol (Welfare Quality®, 2009). This protocol is suitable for assessing on-farm AW by measuring welfare measures to obtain an overall welfare score. However, this protocol is primarily based on animal-based measures what makes it difficult to draw firm conclusions about the systems types in general, as it also assesses the AW impact of farmer management instead of only the characteristic system attributes. On top of that, Welfare Quality (WQ) does not use a zero value and golden standard which makes it difficult to interpret these scores. A way to compare different production systems would be to calculate the differences in scores relative to the conventional system (Gocsik *et al.*, 2016).

Nowadays, most of the pigs are housed in a conventional system as adjustments for higher animal welfare usually result in higher production costs (Gocsik *et al.*, 2015). The farmers base their choice for a certain production system mostly on financial factors as viability, income, and risk because the pig sector is mainly cost-driven (Gocsik *et al.*, 2014). The choice for a system beyond conventional is a completely voluntary choice as it is above legal standards. However, because of the public concern, farmers are increasingly expected to produce beyond the legal minimum standards. In order to keep their license to produce, it is important for the sector to respond to these concerns (Backus and Schans, 2000).

Investing in a more animal-friendly system has an impact on production costs and AW level. For farmers, it is valuable to know which system attributes contribute to what extent to a better level of AW on their farm. The ratio between gain in AW and increase in production costs will give the cost-efficiency. This cost-efficiency is a valuable information source for farmers and other stakeholders when it comes to the decision-making for a certain system type. For broilers, this cost-efficiency is already calculated but in the pig sector, it is up to now unknown (Gocsik *et al.*, 2016).

This research has the objective to gain insight in the cost-efficiency of animal welfare improvements resulting from the conversion conventional to a different pig production system in the Dutch growing-finishing pig sector. Based on the research objective, several sub questions follow:

1. *Which Dutch growing-finishing pig systems exist and which are suitable for this analysis?*
2. *What is the attributional WQ score of each selected system and the individual measures?*
 - a. *Which system attributes are contributing to AW?*
 - b. *To which extent are these attributes contributing to animal welfare?*
3. *What are the production costs of the growing-finishing pig production system types?*
4. *What is the cost-efficiency of the different pig production system types and the individual measures?*

2 Methodology

2.1 General approach

In order to gain insight in the cost-efficiency of animal welfare improvements, a four-step approach was developed. This approach is derived from the method of Gocsik *et al.* (2016) because the Dutch poultry and finishing pig sector are quite similar, as they share the same cost structure and are both economies of scale. The four-step approach is schematically displayed in Figure 2. Each sub-question mentioned in the introduction is linked to a step in the scheme. After a short introduction of the steps, they will be described in more detail in this section.

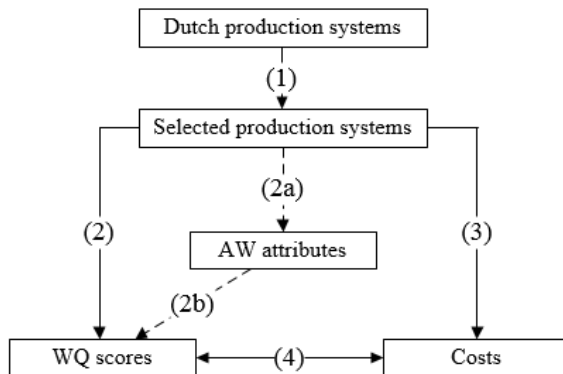


Figure 2: Overview of the research steps

The first step in the research was an inventory of the different pig production systems in the Netherlands. On the basis of this overview, a selection was made with the most suitable systems for this analysis. This selection was based on literature, assessment of an expert and availability of data.

In order to calculate the cost-efficiency, the production costs and the welfare scores of the production systems were calculated. The process of calculating the welfare level of the systems (step 2 in Figure 2) was done with the Welfare Quality® Assessment protocol for pigs (Welfare Quality®, 2009). Because the protocol uses mostly animal-based measures to assess the welfare level on a farm, two sub-steps were necessary to obtain a general system score. In step 2a the selected systems were described according to their characteristic AW system attributes. These attributes have an effect on different aspects of the welfare of the pig and will influence the WQ score. In step 2b the AW attributes were linked to the welfare measures of the protocol based on literature review and the consulting of experts. After that, weights were assigned in a stepwise manner to each link between a system attributes and a welfare measure, according to their relative importance (Gocsik *et al.*, 2016). With these weights and data from the research of (Vermeer, van Reenen and Spoolder, 2012) the WQ index score could be calculated for each system.

The production costs (step 3 in Figure 2) of the selected systems were calculated with the help of a deterministic model described by Gocsik *et al.* (2015). Some of the numbers used in the model were updated with literature, for example, the KWIN-V 2018-2019, and consulting experts (KWIN, 2018). With these production cost calculations, the additional costs (ΔTC) of the different production systems could be compared to the conventional system. The production costs were further analysed by comparing the fixed and the variable costs.

The final step (step 4 in Figure 2) was to link the welfare scores and production costs in order to determine which system is most cost-efficient. The cost-efficiency ($\Delta WQ / \Delta TC$) ratio was calculated for each system compared to the conventional system. This step includes a sensitivity analysis in order to test the strength of the conclusions and to check the influence of price fluctuations.

2.2 Step 1: Selection of suitable production systems

At first, an overview of the current existing pig production systems in the Netherlands was obtained. As mentioned before, in the Dutch pig sector a conventional-, middle- and top-segment can be distinguished regarding AW. The conventional system has to produce according to the minimum legal standards by the Dutch government (European Union, 1998; Wet Dieren, 2015). The middle and top segment produce beyond the legal requirements and can be visualised using the Beter Leven quality mark (Beter Leven Keurmerk, 2019a). This quality mark works with a one- to three-star system. The middle segment consists of systems with one star and the top-segment with two or three stars. The conventional system receives no star. The product receives respectively one, two or three stars when they meet the requirements of the quality mark. The requirements are based on multiple system attributes that contribute to better welfare, for example, square meters per animal, outside access, bedding and enrichment. The specific requirements can be found online (Beter Leven Keurmerk, 2019b).

The Beter Leven quality mark will be used to select different production systems. Ing. R. Hoste was consulted to obtain an overview of the existing concepts and their rank in the star system. Table 1 shows the main concepts in the Netherlands, however, more small concepts exist. The concepts are ordered within the star system, only not all of the concepts are officially certified with the quality mark.

Table 1: Overview of the main existing pig production systems in the Netherlands, ranked according to the Beter Leven quality mark system.

Conventional	Beter Leven 1-star	Beter Leven 2-star	Beter Leven 3-star
Conventional	Varken van Morgen Frievar Wroetvarken Vion Good Farming Star Keten duurzaam Varkensvlees	Free-range Hamletz Livar Piggy's Palace	Organic De Groene Weg

Out of these systems, a selection was made based on available data, literature and expert opinion. Table 2 gives an overview of the selected systems and is based on the table from Gocsik *et al.* (2015). In Beter Leven 1-star, two sub-categories were made with small and large groups because there are some differences. Because the groups are bigger (>40 instead of 8 to 20) and therefore the pens larger, the square meter per pig for large groups are allowed to be minimum 0.9 instead of 1 (Beter Leven Keurmerk, 2019b).

The 2-star systems are rare in the Netherlands. Most of the 2-star meat in Dutch supermarkets is imported from England (Vermeer, 2019). Nevertheless, it was decided to include this system in the analysis as it gives a more complete image of the costs-efficiency of AW improvements. All the farms that are organically certified according to the SKAL certification are automatically classified as 3-star Beter Leven (Beter Leven Keurmerk, 2019b).

In contrast to broiler production, there is no clear difference in the breed of pig that is used in the systems (Gocsik *et al.*, 2016; Vermeer, 2018). Therefore the type of pig is assumed to be equal and not taken into account in the variables of the systems.

As visible in Table 2, the main difference between the systems are found in indoor- and outdoor space, bedding, group size, daylight, enrichment, castration and tail docking (Gocsik *et al.*, 2015). The system characteristics mentioned in Table 2 match the requirements used by the Beter Leven quality mark.

Table 2: Overview of the selected Dutch pig production system types and their characteristics regarding AW

Variable	Production system				
	Conventional	Middle-market segment		Top-market segment	
	Conventional	Beter Leven 1 star small groups	Beter Leven 1 star large groups	Free-range	Organic
Stars 'Beter Leven'	0	1	1	2	3
Indoor space (m ² /110 kg finishing pig)	0.8	1	0.9	0.7	1.3
Outdoor space (m ² /110 kg finishing pig)	0	0	0	0.7	1
Solid floor (%)	40	40	40	100	50
Bedding	Concrete, litter	Concrete, litter	Concrete, litter	Concrete, straw	Concrete, straw
Group size (# pigs per group)	8 to 20	8 to 20	> 40	8 to 30	8 to 30
Daylight in the stable	no	no	no	yes	yes
Enrichment	Metal chain with ball	Wood, sturdy rope, straw, and special scrub	Sawdust, and special scrub	Straw, roughage, and special scrub	Straw, roughage, and special scrub
Castration	yes	no	no	yes	yes
Tail docking	yes	yes	yes	no	no

2.3 Step 2: Calculation of the welfare quality index score for production systems and individual measures.

The calculation of the AW level of the selected production systems was carried out with the Welfare Quality® Assessment protocol for pigs (Welfare Quality®, 2009). In this section, the protocol will be further explained. Furthermore, step 2a: *Decomposition of production systems into system attributes* will be described. Step 2b: *Assigning WQ scores to the system attributes* will be described in detail by addressing the linkages and weights between system attributes and welfare measures, the data that is used, and the approach of the score calculation.

2.3.1 Welfare Quality® Assessment protocol for pigs

The development of the Welfare Quality® protocol is a project stimulated by the European Commission in order to develop a measure that measures the level of animal welfare systematically. It is based on the widely used five freedoms (Webster, 2001). This protocol is innovative in the way that it uses animal-based measures. An advantage of these animal-based measures is that it shows the outcome of both the effects of housing design and management and their interaction. In this research the Welfare Quality® applied to growing and finishing pigs was used. The protocol consists of four welfare principles: Good feeding, good housing, good health, and appropriate behaviour. These principles are divided into twelve welfare criteria that each consists of their corresponding welfare measures, shown in Figure 3. The data produced by the measures are interpreted and synthesised into criterion score and the criteria scores into principle scores (Figure 4) (Welfare Quality®, 2009). In this section, criteria and their corresponding measures (shown in bold), and the selection of measures for further analysis are explained. In this research only the measures scores were used as they were most suitable for visualising the different effects of the system attributes.

	Welfare criteria		Measures
Good feeding	1	Absence of prolonged hunger	Body condition score
	2	Absence of prolonged thirst	Water supply
Good housing	3	Comfort around resting	Bursitis, absence of manure on the body
	4	Thermal comfort	Shivering, panting, huddling
	5	Ease of movement	Space allowance
Good health	6	Absence of injuries	Lameness, wounds on the body, tail biting
	7	Absence of disease	Mortality, coughing, sneezing, pumping, twisted snouts, rectal prolapse, scouring, skin condition, ruptures and hernias
	8	Absence of pain induced by management procedures	Castration, tail docking
Appropriate behaviour	9	Expression of social behaviours	Social behaviour
	10	Expression of other behaviours	Exploratory behaviour
	11	Good human–animal relationship	Fear of humans
	12	Positive emotional state	Qualitative Behaviour Assessment (QBA)

Figure 3: The principles, criteria and measures of the WQ protocol for pigs

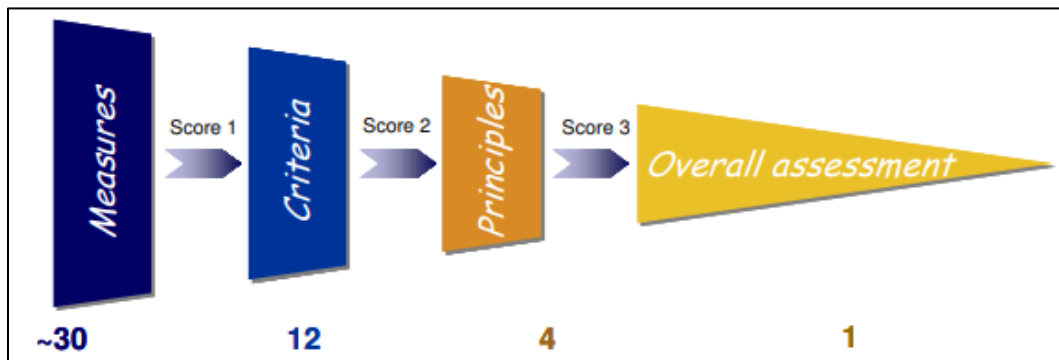


Figure 4: The approach of WQ for integrating the data on the different measures to an overall assessment of the animal unit.

Criterion 1 and 2: Absence of prolonged hunger and absence of prolonged thirst

Welfare quality protocol measures **water supply** available per pig and the **body condition score** in order to score these criteria. Dutch farmers are required to provide their animals 24 hour access to water (Wet Dieren, 2015). Besides that, in the growing/finishing pig industry, it is common that the pigs have non-stop access to food (Vermeer, 2018). Therefore, it is not expected that large differences exist between the production systems. If there are differences, it has to do with farmer management instead of housing design. Therefore, it is decided not to include these measures into the system analysis.

Criterion 3: Comfort around resting

There are two measures belonging to the criterion comfort around resting: **bursitis** and **manure on body**. The score for bursitis reflects the comfort of the lying material in the stable as it is the result of a pressure injury on the weight-bearing points on the legs (Welfare Quality®, 2009). Manure on body is scored in order to assess the cleanliness of the pen and the lying areas. As the type of bedding differs among the systems, these measures are included in the analysis.

Criterion 4: Thermal comfort

The scores for **shivering**, **panting** and **huddling** reflect the thermal comfort of the pigs. The behaviours shivering and huddling are related to temperatures that lay below the thermoneutral zone of the pig and express hypothermia (Miao, Glatz and Ru, 2004). Panting is displayed when pigs are overheated and exposed to temperatures above their thermoneutral zone (Widowski, 2010). Multiple system attributes are influencing these scores and panting and huddling are therefore included in the analysis. Shivering is almost never displayed and was left out (Vermeer, 2018).

Criterion 5: Ease of movement

The only measure that is used to indicate ease of movement is space allowance. This space allowance is scored by measuring the **stocking density** of the stables. The stocking density resembles how much space the animals have to move around. This measure is a housing based measure and is, therefore, an important measure to include in the analysis.

Criterion 6: Absence of injuries

The criterion of absence of injuries is expressed with the measures: **lameness**, **wounds on body** and **tail biting**. Injuries influence welfare directly due to pain and indirectly due to limitations in expressing some behaviours (Brooshooft *et al.*, 2014). Injuries can originate from environmental factors, like housing conditions, or they can be caused by (aggressive) behaviour of pen mates (Quiniou *et al.*, 2010). All three types of injuries are quite common in all systems but the extent varies between the production types. System attributes have an influence on the environmental conditions and indirectly on the behaviour of the pigs. All three measures are therefore included in the analysis.

Criterion 7: Absence of disease

This criterion consists of thirteen measures. The most extreme form of the display of the presence of disease in animals is death. Therefore, **mortality** is one of the measures of absence of disease. However not all mortality is due to disease and mortality itself is not necessarily the main welfare problem. More important is the suffering before the animal his death, and disease often causes suffering (Broom, 1991). For this reason and the fact that mortality is rather low in the finishing pig sector, it was decided to leave this measure out of the analysis.

Respiratory disorders are covered with the measures: **coughing**, **sneezing** and **pumping**. Frequently coughing, sneezing or pumping are clear symptoms of the presence of a respiratory disease (Silva *et al.*, 2009). However, these symptoms can be caused by a large variety of factors and diseases. Because of the number of causes, it was not possible to give a clear, literature-based, description of the link with the system attributes. It was expected that slaughterhouse data on pneumonia and pleurisy are sufficient to cover the differences in respiratory diseases between the systems (Vermeer, 2018). Coughing, sneezing, and pumping are therefore not included in the analysis.

Twisted snouts is another measure of absence of disease. Twisted snouts are characteristic of atrophic rhinitis and can vary from a slight deform of the snout to a more severe nasal distortion (Welfare Quality®, 2009). This disease used to be a widespread disease among the swine population with for example a 40-90% herd incidence in the U.S. (Hogg and Switzer, 1978). Nowadays the prevalence of atrophic rhinitis and therefore, twisted snouts, is very low (Vermeer, 2018). Because of the low prevalence, it was decided to leave this measure out of the analysis.

Health problems in the area of enteric disorders are assessed with two measures: **rectal prolapse** and **scouring** (Welfare Quality®, 2009). Rectal prolapse (when the rectum is pushed out of the anus) and scouring, otherwise known as diarrhoea, are signs of problems in the digestive system. The measure rectal prolapse was not included in the analysis due to the lack of literature about the link with system attributes.

An overall measure for assessing the overall health of the pig is by checking its **skin condition**. Certain diseases can cause characteristic inflammation or discolouration of the skin (Welfare Quality®, 2009). However, it is more a measure that indicates if the pig is feeling well or not and is too indistinct to find linkages in literature. Therefore skin condition was not included in the research.

The last measure that is recorded on the farm is **ruptures and hernias**. A rupture and/or hernia is the protrusion of a body structure or organ through a body wall that would normally contain it, and it results in a lump under the skin. These can be in the umbilical or inguinal area (Welfare Quality®, 2009). These hernias are quite common in current pig production systems (Petersen, 2008). There are strong indications that these conditions are painful to pigs and therefore reduce their welfare (Schild *et al.*, 2015).

Besides these on-farm measurements, slaughterhouse data was collected and form four more measures for the criterion absence of disease. As earlier mentioned slaughter checks were done for **pneumonia** and **pleurisy**. The lungs of the pig are checked for signs of these diseases. Besides the lungs, the heart is checked for **pericarditis** and there is a check for **white spots on the liver**. These white spots are caused by worms that the pigs get inside their system by contact with faecal matter (Roepstorff and Nansen, 1998). Even though there is a difference in infection risk between the systems, due to deworming, in the end, there is no real difference between the systems when it comes to white spots on the liver (Vermeer, 2018). For pericarditis, there was no clear link with a system attribute found in literature. As Alban, Petersen and Busch (2015) state: There is no obvious explanation for the higher occurrence of pericarditis in organic/free-range pigs. Pericarditis and white spots on the liver were excluded from the analysis.

Criterion 8: Absence of pain induced by management procedures

The management procedures this criterion is referring to are **castration** and **tail docking**. These procedures cause pain for the pig. Therefore, it is assessed if these procedures are implemented on the farm or not and if they are if they are performed with or without anaesthesia. Castration of piglets without anaesthesia is not allowed in the Netherlands (Backus, 2004). In some of the Beter Leven star systems the performance of castration and tail docking is not allowed and therefore the AW level of this criterion differs among the systems and was included in the analysis (Beter Leven Keurmerk, 2019b).

Criterion 9: Expression of social behaviours

Pigs are social animals and frequently interact with each other. The behaviour towards another individual can be either negative or positive. **Negative social behaviour**, including biting, reduces the welfare of the inflicted pig. The proportion of negative behaviour with respect to total social behaviour shows in which way the pigs are expressing social behaviour. Many system attributes influence the amount of (negative) social behaviour expressed by the pigs (Ewbank and Bryant, 1972).

Criterion 10: Expression of other behaviours

The measure that displays the expression of other behaviours is the exploratory behaviour. The behaviours that are recorded are **investigation of the pen** and **exploring enrichment material**. Investigation of the pen is defined as sniffing, nosing, licking or chewing all features of the pen. Exploring enrichment material is defined as play/investigation towards straw or other enrichment material (Welfare Quality®, 2009). The design of the pen and the type of present enrichment material defer amongst the production systems and therefore a difference is expected in the expression of other behaviours.

Criterion 11: Good human-animal relationship

The **fear of humans** is assessed by checking whether the animals show a panic response towards humans or not. Although the outcome of this measure is largely defined by the management of the farmer, it is expected that system attributes also influence this measure.

Criterion 12: Positive emotional state

This last criterion is measured by the **Qualitative Behaviour Assessment** (QBA). This assessment considers how animals behave and interact with each other and the environment (Welfare Quality®, 2009). This measure makes an effort to measure the positive state of an animal when an animal is feeling well. Nevertheless, literature describes no distinct factor that influences this measure. Therefore it is hard to link them to system attributes and it was left out the analysis.

To conclude, there were 18 measures selected as WQ measures used in this study. Table 3 gives an overview of these selected measures and their corresponding criteria and principles.

Table 3: Display of the selected WQ measures and the corresponding criteria and principles

Welfare principles	Welfare criteria	Selected welfare measures
Good Housing	3. Comfort around resting	1. Bursitis 2. Manure on body
	4. Thermal comfort	3. Panting 4. Huddling
	5. Ease of movement	5. Space allowance
Good health	6. Absence of injuries	6. Lameness 7. Wounds on body 8. Tail biting
	7. Absence of disease	9. Scouring 10. Ruptures and hernias 11. Pneumonia 12. Pleurisy
	8. Absence of pain induced by management procedures	13. Castration 14. Tail docking
Appropriate behaviour	9. Expression of social behaviours	15. Negative social behaviour
	10. Expression of other behaviours	16. Investigation of the pen 17. Exploring enrichment material
	11. Good human-animal relationship	18. Fear of humans

2.3.2 Decomposition of production systems into system attributes.

As shown in Table 2, there are certain system attributes that defer amongst the different production systems. Some of the variables in the table were taken directly as attributes for the analysis, these were: bedding, group size, enrichment, castration, and tail docking. The variables indoor space and outdoor space were changed to stocking density (total space in m² inside and outside) and outdoor access. In this way, the effects for space per pig and exposure to outside conditions can be assessed separately. Solid floor percentage was left out as fully slatted floors are not allowed in the Netherlands and most of the difference in floor type will be covered by the attribute bedding (Vermeer, 2018). No literature was found on the effect of natural daylight instead of artificial light for pigs, therefore this variable was not considered in the analysis.

The following system attributes were used in this study:

- A1. Stocking density
- A2. Outdoor access
- A3. Bedding
- A4. Group size
- A5. Enrichment
- A6. Castration
- A7. Tail docking

2.3.3 Linkages and weights between system attributes and welfare measures

For the reason that the scores for the WQ measures and the system attributes are not one to one comparable, a method had to be developed to link these two. This method is already described in Gocsik *et al.* (2016) and was adjusted to match the current analysis. Data was available on the measure scores of 79 Dutch farms. In order to transform these measure scores to scores per system attribute, linkages between these two were established based on literature review and the expert opinion of ir. HM Vermeer. These linkages are visualised in Table 4. If there is a X visible in the matrix, a link exists and the system attribute has an effect on that specific WQ measure. If a smaller x is visible, a weak link exists and the shaded cells were based on expert opinion.

The relative importance of the established linkages is shown in Table 5. The weights were necessary to account for the difference in importance between the linkages as some measure have a larger influence than others. If literature or expert opinion gave no reason to assume a difference, the weights were set equal. The weighting procedure was executed as follows: A matrix was made with the relative numbers of importance with the help of ir. HM Vermeer. Each time all the links belonging to one specific measure were compared. When a link had the most influence it received the highest number, when less important a lower number. When they are equally important, they receive the same number. A two is twice as important as a one.

When this importance matrix was established, the weights were calculated in order to be able to make use of allocation formulas. The weights were calculated by dividing 1 by the total sum of the links for a certain measure and then multiply it with the number of the specific link. For example for measure 2. manure on body: stocking density is most important (2), after that bedding and group size are both less important (both a 1). For the weight this means: 1 divided by (2+1+1) =0.25. Stocking density receives a weight of $2 \times 0.25 = 0.5$, group size receives $1 \times 0.25 = 0.25$ and bedding receives $1 \times 0.25 = 0.25$.

Table 4: Matrix showing the linkages between the selected Welfare Quality measures and the selected system attributes

WQ measures	System Attributes						
	A1. Stocking density	A2. Outdoor access	A3. Bedding	A4. Group size	A5. Enrichment	A6. Castration	A7. Tail docking
1. Bursitis	X	X	X			X	
2. Manure on body	X		X	x			
3. Panting		x	x				
4. Huddling		x					
5. Space allowance	X						
6. Lameness	X	x	X	X		X	
7. Wounds on body	X	X	X	X	X	X	X
8. Tail biting	X	X	X	x	X		X
9. Scouring							X
10. Ruptures and hernias		X					
11. Pneumonia	X	X	x	X			
12. Pleurisy	X	X	x	X			
13. Castration						X	
14. Tail docking							X
15. Negative social behaviour	X		X	X	X	X	
16. Investigation of the pen	x	x	X	x	X		
17. Exploring enrichment material	x		X	x	X		
18. Fear of humans			X	X	x		

Legend:	
X	Linked
x	Weak link
	Expert opinion

Table 5: Matrix showing the assigned weights to the linkages between the WQ measures and the system attributes

WQ measures	System Attributes						
	A1. Stocking density	A2. Outdoor access	A3. Bedding	A4. Group size	A5. Enrichment	A6. Castration	A7. Tail docking
1. Bursitis	0.18	0.09	0.55			0.18	
2. Manure on body	0.5		0.25	0.25			
3. Panting		0.5	0.5				
4. Huddling		1					
5. Space allowance	1						
6. Lameness	0.29	0.14	0.14	0.29		0.14	
7. Wounds on body	0.14	0.14	0.14	0.14	0.14	0.14	0.14
8. Tail biting	0.14	0.14	0.21	0.07	0.21		0.21
9. Scouring							1
10. Ruptures and hernias		1					
11. Pneumonia	0.4	0.2	0.2	0.2			
12. Pleurisy	0.4	0.2	0.2	0.2			
13. Castration						1	
14. Tail docking							1
15. Negative social behaviour	0.3		0.2	0.1	0.2	0.2	
16. Investigation of the pen	0.14	0.14	0.29	0.14	0.29		
17. Exploring enrichment material	0.13		0.25	0.13	0.50		
18. Fear of humans			0.33	0.33	0.33		
Total	4.12	3.56	3.76	1.85	1.68	1.67	2.36

In this method of establishing the links and weights, any interactions between factors were not considered resulting in a solely additive method. However, these interactions between the effect of the system attributes (environmental factors), genetic or management factors do exist (Broom and Fraser, 2015). Nevertheless, literature about these interactions is very scarce and could therefore not be included in the allocation formulas. Welfare Quality also states that the measures are unrelated even though interactions exists.

The linking of the system attributes to the WQ measures is a crucial step in obtaining a system welfare score. Nevertheless, no scientific based method was available and for that reason, the same method as in Gocsik *et al.* (2016) was used to establish the weights and linkages. It is realized that this method has some risk of subjectivity. Therefore, the links have a thorough foundation in scientific literature and were the basis for the sensitivity analysis.

In the following part, each link will be further explained by addressing the corresponding literature and expert opinion. They will be described measure-wise.

1. Bursitis

The WQ measure Bursitis was linked to four system attributes: stocking density, outdoor access, bedding, and castration. As earlier mentioned, bursitis is the result of a pressure injury on the weight-bearing points on the legs. When pigs have to rest on a hard concrete- or slatted floor the risk on bursitis is higher (Gillman *et al.*, 2008). Softer bedding like straw reduces the prevalence and severity of bursitis (Smith, 1993; Jensen, 2009; van de Weerd and Day, 2009; Temple *et al.*, 2012). High stocking density also tends to increase the risk and severity of bursal lesions as less comfortable lying area is available (Smith, 1993), especially when pigs become bigger and are forced to lie on a less comfortable areas like the slatted floor (Vermeer 2018). Outdoor access can reduce the prevalence of bursitis when there is a comfortable resting surface, like soil, present outside (Avé *et al.* 2012; Guy *et al.* 2002; Park, Min and Oh 2017). Nevertheless, almost all the outside facilities for fattening pigs in the Netherlands consist of slatted or concrete floors. There is also confounding as systems with outdoor access often have straw bedding indoors (Vermeer 2018). Castration influences the risk on bursitis in the way that castrated pigs are less active and tend to lie more. Therefore barrows have a higher risk at bursal lesions (Quiniou *et al.*, 2010). Literature on group size showed no difference between the prevalence of bursitis (Meyer-Hamme, Lambertz and Gauly, 2016).

Bedding was indicated as the most important factor. Stocking density and castration both equally important and outdoor access less important as part of the difference is probably found because of the straw bedding that is present indoors in free-range systems. The linkages received the following weights: Bedding 0.55, stocking density and castration 0.18 and outdoor access 0.09.

2. Manure on body

In literature was found that stocking density, bedding, and group size are linked to the measure manure on body. The bodies of pigs were dirtier in intensive systems than in extensive systems. The reasoning behind it is that pigs with more space could better separate the defecating and lying regions (Fu *et al.*, 2016). Pigs that are housed on straw have a poorer hygiene score and have more manure on their body (Scott *et al.*, 2006; Jensen *et al.*, 2012; Temple *et al.*, 2012). This is due to the fact that, the manure sticks on the straw and the straw makes the drainage through the slatted floor more difficult. Pigs that are housed in large groups tend to have more manure on their body (Meyer-Hamme, Lambertz and Gauly, 2016). With increasing group size it's more difficult for the pigs to find a dry lying area. Often the percentage of solid floor is less in large groups, or the solid floor is dirty because dunging behaviour is much more diffuse than in smaller pens (Vermeer, 2018). However, this is only a weak link as in most cases there is enough space for the pigs to prevent these problems.

Because the effect of group size is indirectly related to stocking density, stocking density was considered the most important link. Bedding only has an effect if the pigs are not able to separate their dunging area from the rest of the pen and is therefore also considered as a less important factor. Stocking density received a weight of 0.5, the other two links received a weight of 0.25.

3. Panting

The prevalence of panting is relatively low in many of the literature studies (Temple *et al.*, 2012). Even though, there are some system attributes that increases the risk of panting: outdoor access and bedding. When there is outdoor access animals have more options for cooling and there will be less heat stress (Spoolder, 2007; Vermeer, 2018). Bedding can have an effect on panting as straw retains heat (Vermeer, 2018).

Both linkages were considered weak links and, received an equal weight of 0.5 as there is no reason to assume they differ in importance.

4. Huddling

Only the attribute outdoor access was linked to huddling. When it is cold outside, pigs with outdoor access have the option to stay inside. Nevertheless, the portal to the outside area could cause a cold draft and the pigs have to huddle to stay warm (Miao, Glatz and Ru 2004; Vermeer, 2018). This link is considered as weak because it is only occurs in some cases. Because there is only one link, outdoor access receives a weight of 1.

5. Space allowance

Space allowance is directly linked to the system attribute stocking density. When there is a lower stocking density, there is more space per pig to move around so more space allowance. Because there is only one link, stocking density receives a weight of 1.

6. Lameness

The measure lameness was linked to five system attributes: stocking density, outdoor access, bedding, group size, and castration. A lower stocking density decreases the risk of injuries to the leg and claw because pigs can easier escape from aggressive interactions (Jensen, 2009; Quinn, 2014). Outdoor access can affect limb health both positively and negatively. A wet and dirty outdoor area can easily spread coronary band infections. On the other hand, when there is outdoor access the indoor area is often cleaner, which is positive for the claws (Vermeer, 2018). Also in literature, you find different outcomes on this matter as also described by Etterlin *et al.* (2014), Sundrum and Weißmann (2005) and Gillman *et al.* (2008). As there is division about the effect, outdoor access is considered a weak link. Furthermore, bedding has an influence on lameness as multiple floor features play a role, such as the proportion of slatted and flooring material (soil, concrete, metal, plastic, rubber), bedding provision/quantity and cleanliness (Willgert *et al.*, 2013). Stalled floors increase the risk of lameness and straw can protect pigs from injuries to the leg and claw (Scott *et al.*, 2006; Jensen, 2009; Lensink *et al.*, 2013). In pens with larger group size, pigs tend to have a worse lameness score due to fighting (Street and Gonyou, 2008; Willgert *et al.*, 2013). The same goes for castration as non-castrated pigs tend to be more aggressive and display mounting behaviour which can result in leg injuries for pen-mates (Fredriksen *et al.*, 2008; Quiniou *et al.*, 2010).

The attributes stocking density and group size were pointed out as most important and received a weight of 0.29. The other attributes were seen as less important because of a low prevalence or both positive and negative effects. They all received a weight of 0.14.

7. Wounds on body

The measure wounds on body was linked to all seven system attributes. Most of the attributes influence the level of aggression. Fighting amongst pigs often inflicts wounds on the body. The linked attributes are stocking density, outdoor access, group size, enrichment, castration, and tail docking. A high stocking density leads to more difficulty to escape from an aggressor and sometimes to more frustration and therefore aggressive behaviour (Turner *et al.*, 2000; Fu *et al.*, 2016). Outdoor access also has a positive influence on the number of wounds as there is more space to flee (J. . Guy *et al.*, 2002). There are also more body lesions found in larger group sizes, which is mostly caused by more aggression at the food trough (Spoolder, Edwards and Corning, 1999; Wolter *et al.*, 2001; Meyer-Hamme, Lambertz and Gauly, 2016). More and/or better quality enrichment improves the number of wounds on the body (Cornale, Macchi, Miretti, Renna, Lussiana, Perona, Battaglini, *et al.*, 2015). Pigs which do not have enriching stimuli to manipulate spend longer periods manipulating pen mates (Beattie *et al.*, 2001). Whether or not castration is performed in the system also has an effect on the number of wounds. Non-castrated pigs are generally more aggressive and have therefore a higher lesion score (Quiniou *et al.*, 2010). The performance of tail docking is also influencing wounds on body as the urge of tail biting is redirected to other body parts (Simonsen, 1995; Nannoni *et al.*, 2014).

The attribute bedding influences the number of wounds on body in another way. Slatted floors are often slippery and may cause pigs to fall. Hence, slatted floors increase the risk of wounds on body with respect to pens with deep bedding (Schröder-Petersen and Simonsen, 2001; Courboulay and Delarue, 2009).

There is no reason to assume any difference in importance between the attributes. Therefore it is decided to give them all a weight of 0.14.

8. Tail biting

Tail biting is linked to six of the seven system attributes, namely: stocking density, outdoor access, bedding, group size, enrichment, and tail docking. Walker and Bilkei (2006) made an overview of the existing literature and most literature concludes that tail docking is more observed when stocking density is high. Stocking density is important in the establishment of a social hierarchy that is linked to the incidence of tail biting. When pigs have outdoor access, tail biting is reduced as outdoor-kept pigs have less social discomfort, more space allowance, and more objects to chew on (Walker and Bilkei, 2006). Guy et al. (2002) also stated that pigs reared in outdoor paddocks were found to spend more time exploring and moving, and less time tail-biting compared to those in an intensive indoor system. Bedding is also an important factor in the incidence of tail biting because when pigs are housed on floors without bedding, they tend to spend more time tail biting (Schröder-Petersen and Simonsen, 2001). Especially straw, hay and peat as manipulable material reduces the probability of tail biting (AHAW European Commission 2014). When it comes to the effect of group size, literature is divided. Some research states that group size did not affect tail biting and stocking density is far more important, while others state that there is a clear effect (Moinard *et al.*, 2003; Schmolke, Li and Gonyou, 2003; Walker and Bilkei, 2006; Vermeer, de Greef and Houwers, 2014). Eventually, with expert opinion, it was decided that group size would receive a weak link. Enrichment material is also a way to reduce tail biting as it keeps the pigs occupied. Straw reduces tail biting but also substrate or feed dispensers are a good alternative. A metal chain (used in conventional systems) is not sufficient for reducing tail biting (Petersen, Simonsen and Lawson, 1995; Beattie *et al.*, 2001; Weerd *et al.*, 2006; Ernst *et al.*, 2018). In literature, no significant difference for tail biting between castrated and non-castrated pigs was shown (Quiniou *et al.*, 2010). The last attribute that has an effect on tail biting is tail docking. Tail docking is seen as an effective way to reduce tail biting as it removes the target and with that the temptation to most extent. However, it does not solve the problem entirely (Moinard *et al.*, 2003).

The linkages with bedding, enrichment and tail docking were seen as the most important links as they have the largest effect on tail biting. These three links received a weight of 0.21. Less important were the linkages stocking density and outdoor access with a weight of 0.14. The least important, with a weight of 0.07, was group size as only a weak link was found.

9. Scouring

Only the attribute tail docking was linked to scouring. Tail docking of piglets increases the risk of diarrhoea (Pearce, 1999). Next to that it was found that slatted floors were linked to a higher prevalence of scouring. The production systems in the Netherlands approximately have the same slatted floor ratio and therefore this link was left out. As the only link with tail docking was found, a weight of 1 was assigned to this link.

10. Ruptures and hernias

Not much literature was found on the linkages between hernias and housing factors. Only the link between hernias and outdoor access was established. Organic and free-range pigs had a lower prevalence of hernias compared to the conventional housed pig. Unfortunately, no data was available that could explain this finding (Alban, Petersen and Busch, 2015). The link between received a weight of 1.

11 and 12. Pneumonia and Pleurisy

These welfare measures can be assessed simultaneously as they have many similarities. Pneumonia and pleurisy are linked to the attributes: stocking density, outdoor access, bedding, and group size. A higher stocking density, and often associated higher air space stocking density, makes it easier for diseases to spread in the barn. The concentration of infectious particles in the air is higher in these situations and increases the risk of pneumonia and pleurisy (Maes *et al.*, 1996; Stärk, Pfeiffer and Morris, 1998; Maes *et al.*, 2000, 2001). For the same reason, pneumonia and pleurisy are also less prevalent in systems with outdoor access (von Borell and Sørensen, 2004; Juska, Juskiene and Leikus, 2013; Alban, Petersen and Busch, 2015). When it comes to the link with bedding, literature is divided and therefore it receives only a weak link. On one side deep bedding can insulate the pig from the floor temperatures and reduces the risk of lung problems. On the other hand, there is an increased risk because of the higher dust level bedding creates (Cleveland-Nielsen, Nielsen and Ersbøll, 2002; Scott *et al.*, 2006; Ruis, Pinxterhuis and Vrolijk, 2010; Sanchez-Vazquez, 2013). There is also a higher chance of lung diseases in larger groups as the chance of a sick animal and the spread of the disease is higher (Maes *et al.*, 2000; Cleveland-Nielsen, Nielsen and Ersbøll, 2002).

Stocking density was seen as the most important factor and received a weight of 0.4. The other three links received an equal weight of 0.2.

13. Castration

The WQ measure castration has a direct link with the system attribute castration and received a weight of 1.

14. Tail docking

Like castration, the measure tail docking is directly linked to the system attribute tail docking and received a weight of 1.

15. Negative social behaviour

A lot of research has been done on the factors influencing negative behaviour in pigs. This measure was linked to the attributes: stocking density, bedding, group size, enrichment, and castration. With higher stocking density, pigs show more aggression and there are more attacks instead of pigs avoiding confrontation, especially after mixing of groups (Turner, Horgan and Edwards, 2001; Schmolke, Li and Gonyou, 2003; Vermeer, de Greef and Houwers, 2014; Fu *et al.*, 2016). The use of straw also reduces aggression in pigs and therefore bedding is linked to negative social behaviour. Pigs show more explorative and playing behaviour and less aggression (Tuytens, 2005). A larger group size both has positive and negative effects on negative social behaviour (Turner, Horgan and Edwards, 2001; Andersen *et al.*, 2004; Rodenburg and Koene, 2007; Averós *et al.*, 2010; Meyer-Hamme, Lambertz and Gauly, 2016). Subordinate pigs or victims can be chased by more pigs in larger groups. However, they can also hide in larger groups (and often larger pens) (Vermeer, 2018). Just like described in 8. *tail biting*, enrichment can reduce aggression and therefore the type of enrichment is linked to negative social behaviour (Beattie *et al.*, 2001; Tuytens, 2005; Scott *et al.*, 2006; van de Weerd and Day, 2009). Castration is also linked to negative social behaviour as non-castrated males tend to be more aggressive (Fredriksen *et al.*, 2008). For borrows, for example, it takes less time to organise the hierarchy between pen-mates than for intact males (Quiniou *et al.* 2010).

In the weight distribution, stocking density was seen as the most important factor, after that bedding, enrichment, and castration and the least important was group size. They received weights of 0.3, 0.2 and 0.1 respectively.

16. Investigation of the pen

Investigation of the pen was linked to the five attributes: stocking density, outdoor access, bedding, group size, and enrichment. With less space per animal, pigs display less locomotion and exploration (Beattie, Walker and Sneddon, 1996; Cornale *et al.*, 2015; Weng, Edwards and English, 1998). Especially pigs lower in rank will perform less exploration (Vermeer, 2018). Therefore, stocking density was linked to investigation of the pen but only with a weak link. Also outdoor access offers more space and variation and offers more possibilities for exploration and receives similar to stocking density, a weak link (Guy *et al.*, 2002; Vermeer, 2018). When pigs are offered bedding, they spend more time examining the bedding material rather than other parts of the pen (Guy *et al.*, 2002; Tuytens, 2005). The link with group size is only a weak link as although research showed a connection, this is probably due to the fact that a larger group is housed in a larger pen with more opportunities for exploration (Averós *et al.*, 2010; Meyer-Hamme, Lambertz and Gauly, 2016). But this is caused by the larger pen, not by the larger group. More different types of enrichment could be available in a large pen but not necessarily more enrichment per pig (Vermeer, 2018). Enrichment is important as pigs are occupied with the enrichment material and feel less urge to explore the pen (Petersen, Simonsen and Lawson, 1995; Beattie *et al.*, 2001; Scott *et al.*, 2006). The type of enrichment is determinative for the size of the effect (van de Weerd and Day, 2009).

Enrichment and bedding were rated the most important and received a weight of 0.29. The other three links were seen as less important and received a weight of 0.14.

17. Exploring enrichment material

This measure was linked to the attributes: stocking density, bedding, group size, and enrichment. Stocking density is weakly linked to exploring enrichment material. When there is more floor space then there is more movement in enriched pens (Beattie, Walker and Sneddon, 1996; Cornale *et al.*, 2015). When pigs are offered bedding, they display more explorative and playing behaviour what is also directed to the enrichment material (Tuytens, 2005). There is some confounding present as pens with bedding often also contain better quality or more enrichment (Vermeer, 2018). Group size is also weakly linked to this measure. In larger groups, there is an increase in activity but it defers between studies if there is an increase in exploration found (Averós *et al.*, 2010; Meyer-Hamme, Lambertz and Gauly, 2016). Larger group size automatically means that there is a larger pen. When pigs are offered the same amount of enrichment, pigs in larger pens could be offered a larger variation in types of enrichment (Vermeer, 2018). Finally, there is an obvious link to the type of enrichment and the measure exploring enrichment material. More or better enrichment material makes sure that the pigs rooting behaviour is directed to the enrichment material (Beattie *et al.*, 2001). There is a difference in the attractiveness of the material and therefore the type of enrichment is linked to the time spend on exploring enrichment material (Da Silva, Manteca and Dias, 2016).

Enrichment was the most important attribute and received a weight of 0.5. After that, bedding was seen as most important and received a weight of 0.25. The attributes stocking density and group size were indicated as least important, as it only has an influence in some cases, and received a weight of 0.13.

18. Fear of humans

The measure fear of humans is mostly influenced by the management of the farmer. Besides that, three system attributes are linked to this measure: bedding, group size, and enrichment. Pigs that are housed on straw display fewer panic responses than pigs that are housed on slatted floors (Courboulay and Delarue, 2009). Also, pigs that are housed in larger groups show an improved human-animal relationship (Meyer-Hamme, Lambertz and Gauly, 2016). For both links, no further explanation was given in research. It's generally accepted that pigs that grow up in enriched environments develop better social skills towards pen-mates. Combined with positive experiences with humans, fear for humans may decrease (Vermeer, 2018; Tuytens, 2005). Therefore, the measure fear of humans is weakly linked with enrichment as it is an indirect link.

There is no reason to assume a difference in importance between these measures. Consequently, they all received an equal weight of 0.33.

2.3.4 Data

The data set that was used for calculating the Welfare Quality measure scores originated from the research of Vermeer, Reenen and Spoolder (2012). This data set contained measurements of 79 farms located in the Netherlands. Information was given on all the measurements required by the Welfare Quality® protocols and therefore existed of 69 data points besides the farm ID. The file also contained the slaughter data, data about the observer, principle scores, the assigned WQ category and distinction between organic and conventional farms, of which only the latter was used.

During the formation of the data set, the Beter Leven quality mark was not yet in use and only the distinction between organic and conventional was made. For that reason, additional data was provided with information about the housing conditions of the measured pens on the farms. On each farm, different numbers of pens were described. Data about the floor type, stocking density, type of enrichment, outdoor access, castration, and tail docking were used to manually evaluate each farm on their corresponding ranking in the star system of the Beter Leven quality mark. In total 16 farms were assigned as 3 stars, 21 farms were indicated as 1 star and 42 conventional farms. The number of 2-star farms in the Netherlands is very limited and are not present in this data set. The numbers of the 2-star system will be estimated using the data on the other systems and expert opinion.

2.3.5 Calculation scores

For the calculation of the measure-, attributional- and overall welfare scores the same step-wise method as in Gocsik *et al.* (2016) was used and is shown in Figure 5.

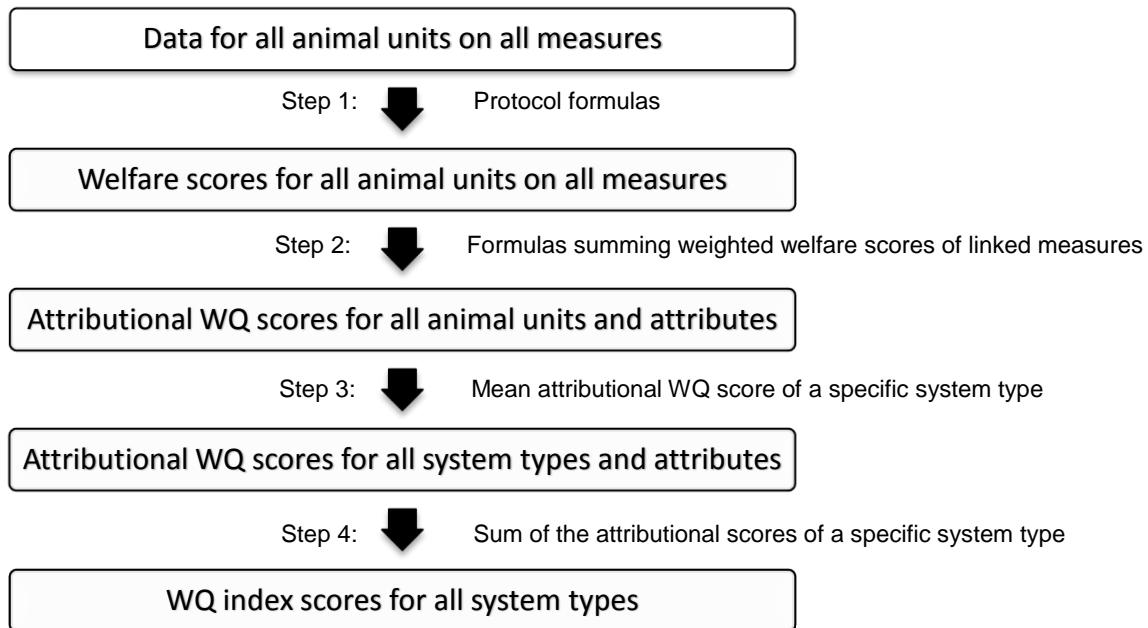


Figure 5: Overview of the WQ score calculations

Step 1: Calculation of WQ measure scores

Firstly, the measure data had to be transformed to the actual WQ measure scores as only the actual recordings were found in the data set. The measure scores were calculated using the Welfare Quality pig protocol and an excel file with parameters from INRA (Welfare Quality®, 2009; INRA, 2018).

Data about the castration and tail docking were transformed to set data per system type. This was done because at the time of data collection the requirements and legislation were not the same as nowadays. For 3 stars, scores were based on the pigs being castrated with anaesthetics and no tail docking was performed, for 1-star scores were calculated for no castration and tail docking without anaesthetics and for conventional castration with anaesthetics and tail docking without anaesthetics.

For the measures pumping, twisted snouts, rectal prolapse and skin condition the % pigs with the condition was calculated using the score 2 based on the interpretation of the protocol. In the protocol, it was not clearly indicated if the calculation used the score 1, 2 or average.

In the protocol, some data points are joint together to obtain a WQ measure score. For thermal comfort a decision tree is used with the scores of shivering, panting and huddling as input to form an overall score for thermal comfort. The same goes for the calculation for the score for pain induced by management that joins castration and tail docking in a decision tree. In this analysis, the measures are assessed individually. For that reason, the scores needed to be split up.

For thermal comfort, huddling and shivering were mutually exclusive so when huddling was recorded, the farm received a score of 34 and when panting was recorded a score of 24. For castration, the sum of the scores for all the options with castration (with anaesthesia) was divided by the sum of the scores without castration. This resulted in a number of 0.833, which means that farms that perform castration receive on average a score that is 0.833 lower than without castration. This was transformed to a measure score of 100 for no castration and 83.3 with castration. The same was done for tail docking which resulted in a score of 100 for systems without tail docking and a score of 39.7 when the tails were docked (without anaesthesia). For the score of absence of disease, all measures are checked if they are on a warning or alarm level. The sum of all the alarms and warnings form a total score for absence of disease. To single out the effect of the measures scouring, ruptures and hernias, pneumonia and pleurisy, a score of 0 was used when these measures had an alarm, 40 when there was a warning and 100 if levels were below the warning threshold. These numbers were derived from the formula used in the protocol. For ruptures and hernias two separate scores were calculated in WQ for severity 1 and 2. The average of these scores was taken for the overall score. Finally the score for the expression of other behaviour needed to be split up into a score for investigation of the pen and exploring enrichment material. Because there was no indication of a best and worst score in the protocol, the lowest measured value in the data set was taken as the 0 point and the highest value as the 100 point.

The measure scores of the non-selected measures were also calculated in order to inspect any large differences between the systems that were not anticipated. No surprising outcomes were found and this data was not used for further analysis.

Step 2: Calculation of attributional WQ scores for all animal units and attributes

In the next step the calculated measure scores for all animal units had to be translated in attributional WQ scores for all attributes. All welfare measures were linked to the system attributes (Table 5) and the attributional WQ for the attributes can be calculated by multiplying each linked welfare score with their assigned weight and summing all the weighted scores that belong to that system attribute:

$$WQ_{A_{jk}} = \sum (w_{ik} * x_{ij})$$

$WQ_{A_{jk}}$ = attributional WQ score for animal unit j ($j = 1...79$) and system attribute k ($k = 1...7$, where 1 = stocking density, ..., 7 = tail docking)

w_{ik} = weight of the link between welfare measure i ($i = 1...18$, where 1 = bursitis, ..., 13 = fear of humans) and system attribute k ,

x_{ij} = the welfare score on welfare measure i and animal unit j

Step 3: Calculation of attributional WQ scores for all system types and attributes

In step 3, scores for all animal units are used to determine the scores for the system types. All animal units are sorted on their Beter Leven star level and mean scores per system type were calculated to obtain the attributional WQ scores per system type.

$$WQ_{A_{km}} = \sum (WQ_{A_{jk}}) / n$$

$WQ_{A_{km}}$ = attributional WQ score for system type m ($m = 1...3$, where 1 = conventional, 2 = 1-star, 3 = organic),

n = number of farms of system type m .

The scores for the free-range system were estimated using the attributional WQ scores for attributes of the other system types. With the expert opinion of ir. HM Vermeer the following estimations were made: The score for stocking density was estimated on the average between 1-star and organic as free-range pigs have less space than organic pigs but more than 1-star. For outdoor access the score of organic was taken as both have outdoor access. The same accounts for bedding, as both free-range and organic uses the same bedding. Group size is roughly the same in all systems so the average of 1-star and organic was used. Enrichment is also better than 1-star and below organic so the average was used. The regulation of castration and tail docking are equal in the free-range system and the organic system so the same numbers were used.

Step 4: Calculation of WQ Index score for all system types

To obtain the WQ index score for all systems, the mean attributional WQ scores for all attributes were summed:

$$WQ_{I_m} = \sum (WQ_{A_{km}})$$

WQ_{I_m} = WQ Index score for system type m .

2.4 Step 3: Calculation of production costs.

2.4.1 Economic model

With the use of the economic model developed by Gocsik *et al.* (2015), the production costs for all the systems were calculated. This model is able to calculate costs (and revenue's) for pig production systems. In the model, the systems conventional, Beter Leven 1-star small groups, Beter Leven 1-star big groups, free-range, and organic were used as they are comparable with the systems in this analysis. The costs were split in fixed and variable costs in order to obtain a more detailed overview and were described per delivered finishing pig.

This research only focusses on the production costs of the systems. This choice has been made as the meat prices can fluctuate severely over time what indicates that there is a degree of risk involved, especially in the top market segment (Agrimatie, 2018). An estimation of the production costs, therefore, contains less uncertainty than an income statement. Nevertheless, the costs could be used as an indication for the breakeven gross revenue per delivered pig.

2.4.2 Data

In addition to the meat price, there are more parameters that vary over time. In the model, several prices and technical variables were used that needed to be updated in order to provide a prediction for the current situation. Some of the parameters were stochastic parameters in the model but were set to deterministic parameters in this study. For the adjustments of these variables and prices for the conventional sector KWIN-V 2018-2019 was consulted (KWIN, 2018). For the variables of the other systems, numbers were adjusted to the change that occurred in the conventional system. Ing. R. Hoste was consulted and delivered feedback. An overview of the variables and prices used in this research is shown in Table 6.

For the prices, an average of the prices for the years 2013-2017 was used based on KWIN-V (KWIN, 2018). Prices for the organic system were calculated using the same price premium as in the study of Gocsik *et al.* (2015). The prices for straw/roughage were retrieved from (Bussink, 2018) and were set at €0.15 per kg. The price for land was set to €65.000 per hectare what is the average of 2017 for livestock farming land in the region south of the Netherlands (ASR real estate, 2019). This region was chosen as it contains the largest share of pig farmers (POV, 2017). Quality discount carcass and costs for transport to slaughterhouse were set at €0. The feed conversion for the organic system was changed back to 3.05 instead of adjusting it to the improvement seen in the conventional system (Hoste, 2019). The investments costs for outdoor access were raised with 10% since the previous number was outdated and the investments costs for buildings in the Beter Leven 1-star small groups and free-range were set equal to the conventional system (Hoste, 2019).

2.5 Step 4: Analyse the cost-efficiency of the pig production system

2.5.1 Calculation of cost-efficiency

The calculated production costs and Welfare index scores per system were combined in order to calculate the cost-efficiency. The conventional system was used as the reference and the other system types were compared with the values of the conventional systems. The change in WQ index score and the change in production costs were calculated with respect to the conventional system. Besides that, changes were calculated for the other systems relative to the system below in the welfare ranking.

2.5.2 Sensitivity analysis

Because there is quite some uncertainty involved in the calculation of the WQ scores (see page 11) and costs results (see page 19), a sensitivity analysis was carried out in order to test the strength of the conclusions. In the process of calculating the WQ Index scores, some arbitrary steps were required. Besides that, management of the farmer also influences the WQ score which causes some possible deviation from the calculated Index score. Therefore, the standard error for the system types was taken to show the range of WQ index scores this data set contains.

On the side of the costs, some uncertainty is involved in the price estimations. The pig meat prices depend on supply and demand and fluctuate constantly. The maximum and minimum price of the years 2013-2017 were taken to visualise the uncertainty around the production costs. For organic, the maximum and minimum price were calculated relative to the change in the conventional prices.

Table 6: Technical inputs, costs and prices used in the economic model

	Unit	Production system				
		Conventional system	Beter Leven 1 star small groups	Beter Leven 1 star big groups	Free Range system	Organic system
Prices						
Piglet price	€/25 kg animal	46.05	46.05	46.05	46.05	115.12
Feed price	€/100kg	24.80	24.80	24.80	24.80	39.68
Technical variables						
Occupation rate	%	94.7	94.7	91.9	94.7	94.7
Mortality rate	%	2.4	2.1	2.4	3.5	4.5
Grow/animal/day	gram	822	853	775	775	758
Feed conversion	kg	2.55	2.55	2.67	2.87	3.05
Start weight piglet	kg	25	25	25	25	25
Finishing live weight	kg	120.6	120.6	120.6	120.6	120.6
Length cleaning/ drying period	days	3	3	3	3	3
Feed intake/animal/day	kg	2.10	2.18	2.07	2.22	2.31
Carcass weight	kg	95.5	95.5	95.5	95.5	95.5
Produced manure dry feed/APFP	ton	1	1	1	1.1	1.2
Used roughage/straw	gram/animal/day	0	15	15	400	400
Labour required	hrs/1.000 pigs/ year	598	659	604	1226	1246
Working hours/FLE	hrs	2349	2349	2349	2349	2349
Stable capacity	# of pigs	4200	4200	4200	4200	4200
Variable costs						
Health care	€/animal	0.80	0.80	0.84	1.09	1.38
Electricity	€/animal	1.00	1.00	1.00	1.00	1.93
Heating	€/animal	0.50	0.50	0.50	0.50	0.46
Water	€/animal	0.70	0.70	0.70	0.70	0.66
Overhead	€/animal	0.60	0.60	0.60	2.16	3.72
Manure disposal	€/ton	19	19	19	19	19
Price straw/roughage	€/kg	0.15	0.15	0.15	0.15	0.15
Price land	€/m ²	6.50	6.50	6.50	6.50	6.50
Certification fee organic/ free-range	€/animal	-	-	-	0.04	0.06
Interest livestock	%	6	6	6	6	6
Fixed costs						
Labour farmer	€/hour	26.90	26.90	26.90	26.90	26.90
Labour hired	€/hour	19.96	19.96	19.96	19.96	19.96
Depreciation buildings	%	4	4	4	4	4
Depreciation inventory	%	8	8	8	8	8
Depreciation outdoor access	%	-	-	-	4	4
Depreciation air scrubber	%	10	10	10	-	-
Maintenance buildings	%	1	1	1	1	1
Maintenance inventory	%	2	2	2	2	2
Maintenance outdoor access	%	-	-	-	1	1
Interest invested capital	%	5	5	5	5	5

3 Results

3.1 Welfare Quality index scores

In the first steps of the calculation of the Welfare Quality index scores, welfare measure scores were calculated for all animal units. The average measure scores of the farms in each system are shown in Table 7. The scores have a range of 0-100 where 0 is the worst possible score and 100 the best score.

Table 7: Mean welfare score per welfare measure and per system type

Welfare measures	System type		
	Conventional	Beter Leven 1 star	Organic
1. Bursitis	93	93	94
2. Manure on body	66	81	89
3. Panting	94	98	95
4. Huddling	100	100	95
5. Space allowance	51	63	82
6. Lameness	95	96	93
7. Wounds on body	64	66	54
8. Tail biting	97	95	85
9. Scouring	93	84	69
10. Ruptures and hernias	95	95	83
11. Pneumonia	19	27	30
12. Pleurisy	91	94	93
13. Castration	83	100	83
14. Tail docking	40	40	100
15. Negative social behaviour	75	79	73
16. Investigation of the pen	38	32	16
17. Exploring enrichment material	7	12	52
18. Fear of humans	71	73	64
Total	1,272	1,329	1,348

Table 7 shows that approximately half of the measures show increasing or stable scores shifting to the higher star systems and the other half shows a decreasing trend. The measures: bursitis, panting, huddling, lameness, pleurisy, negative social behaviour and fear of humans all score roughly the same in the three systems with no more than 7 points difference. The main differences are found in the measures: space allowance (31 points), tail docking (60 points) and exploring enrichment material (45 points) and are in favour of the organic system. The measure pneumonia also shows a smaller increase of 11 points. Nevertheless, the total score only shows small differences as the positive effect of these main effects are compensated by a couple of small negative trends. These smaller negative trends are: tail biting, scouring, ruptures and hernias, and investigation of the pen and show a difference between 24 and 12 points. The measures wounds on body and castration do not show an increasing or decreasing trend as the score of Beter Leven 1 star was the highest score of the three. Appendix I gives an overview of the trends in the calculated measure scores for the systems in the form of boxplots derived with IBN SPSS Statistics 23.

With these measure scores for each system and the assigned linkages and weights, attributional scores for all system attributes for all systems were calculated and are shown in Table 8 and Figure 6. In this table, the estimates scores of the free-range system are included.

Table 8: Mean attributional welfare scores per system attribute and per system type

System attributes	System type			
	Conventional	Beter Leven 1 star	Free Range	Organic
A1. Stocking density	224	250	260	270
A2. Outdoor access	315	318	293	293
A3. Bedding	231	240	238	238
A4. Group size	119	127	126	125
A5. Enrichment	83	85	89	93
A6. Castration	138	156	136	136
A7. Tail docking	163	153	195	195
Total	1272	1329	1336	1348

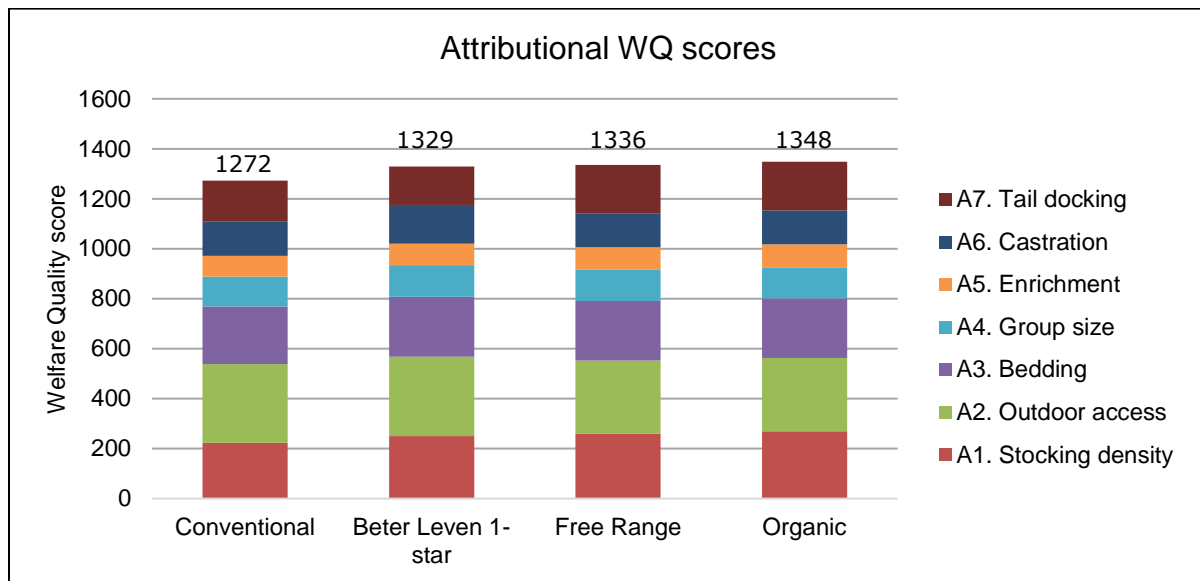


Figure 6: Mean attributional welfare scores per system attribute and per system type

These results show, similar to the measure scores, various trends. Three of the system attributes show the highest scores for the organic system: stocking density, enrichment, and tail docking. This leads back to the trend observed in the welfare measures space allowance, exploring enrichment material and tail docking. These measures scored highest for the organic system and received relatively high weights scores for the linkages with these three system attributes. There is a difference, compared to the measure scores, in the fact that none of the attribute scores was highest in the conventional system. Instead, more attributes scored highest for the 1-star system: outdoor access, bedding, group size, and castration. The score for castration is easily traced back to the fact that the 1-star system is the only system that does not castrate their pig according to the Beter Leven standards. The other three attributes that scored highest for 1-star are attributes that were linked to measures that were relatively complex in their contributing factors.

Where a lower stocking density only showed positive effects in literature, change in outdoor access, bedding and group size show a trade-off between positive and negative effects. Scores for these attributes are composed out of the linked measures scores and their weights. The scores of the measures linked to outdoor access (for example wounds on body, tail biting and ruptures and hernias) were lower for the organic system where outdoor access is provided. This resulted in an overall negative effect of outdoor access on the level of animal welfare. Bedding had a positive effect for the 1-star system compared to the conventional system but this positive effect did not increase further in the free-range and organic system. The positive and negative effects are visualised in Figure 7.

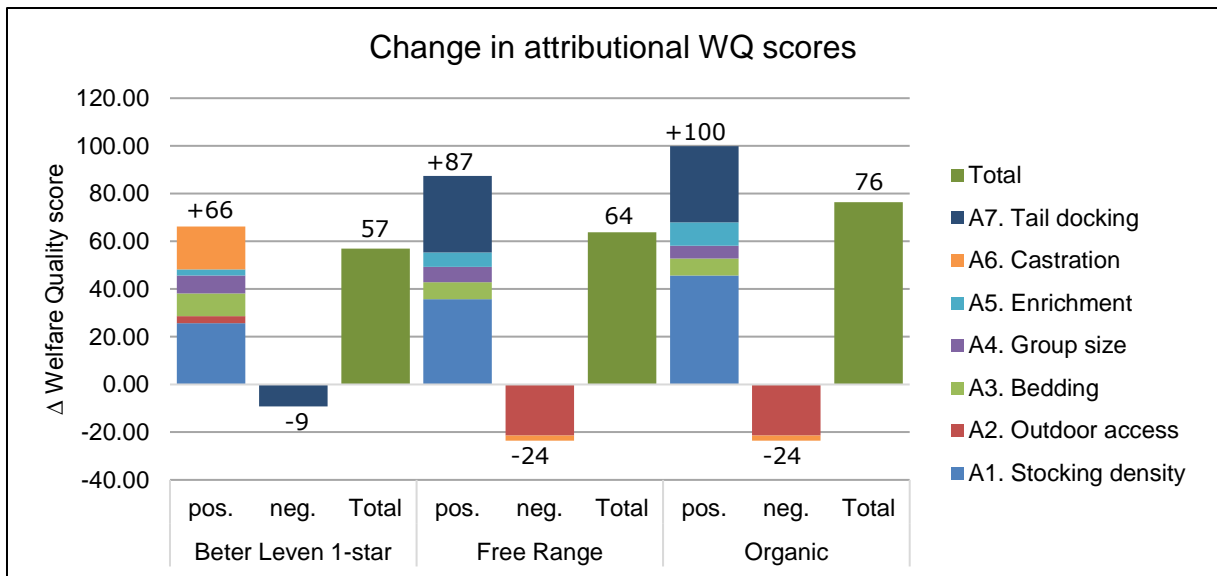


Figure 7: Change in attributional welfare score of the system types relative to the conventional system

Comparing the overall effect of the attributes to the AW level on farms, it is visible that outdoor access has the most effect with 23% of the total score. The second and third most contributing factors were stocking density (19%) and bedding (18%). Thereafter, the attributes tail docking (13%), castration (11%), group size (9%) and enrichment (7%) follow. The cause of the large contribution of outdoor access, stocking density and bedding lies mainly with the weights assigned to the linkages. The attributes were linked to relatively many measures and received higher weights.

The total scores in Table 7 and 8 reflects the overall WQ Index score per system. As already mentioned, only small differences are visible and the difference between the lowest and highest score is only 6.0%. The organic system scores best with a total WQ score of 1348 after that in decreasing order the free-range system with a score of 1336, the 1-star system with a score of 1329 and finally the conventional system which scored 1272.

3.2 Production costs

The variable and fixed costs for all system types, including small and big groups for 1-star, were determined using an updated version of the model of Gocsik *et al.* (2015). The formula's used by the model are given in Appendix II. The results of the variable costs are presented in Table 9 and Figure 8.

The total variable costs show an increasing trend from conventional to the organic system. There is a small increase changing from conventional to Beter Leven 1-star small groups (€0.65) and from small groups to big groups (€2.95). Converting from Beter Leven 1-star big groups to the free-range system shows a larger increase of €18.30 and finally the shift to organic system almost doubled the variable costs per delivered pig with an increase of €121.06.

Table 9: Variable costs (€/delivered pig) per system type

	System type				
	Conventional	Beter Leven 1 star small groups	Beter Leven 1 star big groups	Free Range system	Organic system
Purchase price piglet	46.05	46.05	46.05	46.05	115.12
Feed costs	57.41	57.41	60.11	64.48	110.43
Health care	0.80	0.80	0.84	1.09	1.38
Water	0.70	0.70	0.70	0.70	0.66
Costs enrichment/straw	0.00	0.25	0.27	7.15	7.32
Overhead	0.60	0.60	0.60	2.16	3.72
Mortality	0.02	0.02	0.02	0.03	0.08
Electricity	1.00	1.00	1.00	1.00	1.93
Heating	0.50	0.50	0.50	0.50	0.46
Interest animals	1.46	1.47	1.58	1.72	3.69
Manure disposal costs	6.21	6.21	6.58	7.24	8.07
Labour hired	0.05	0.45	0.00	4.39	4.63
Lease production quota	2.61	2.61	2.77	2.77	2.83
Certification fee organic/ free-range	0.00	0.00	0.00	0.04	0.06
Total variable costs	117.42	118.07	121.02	139.32	260.38

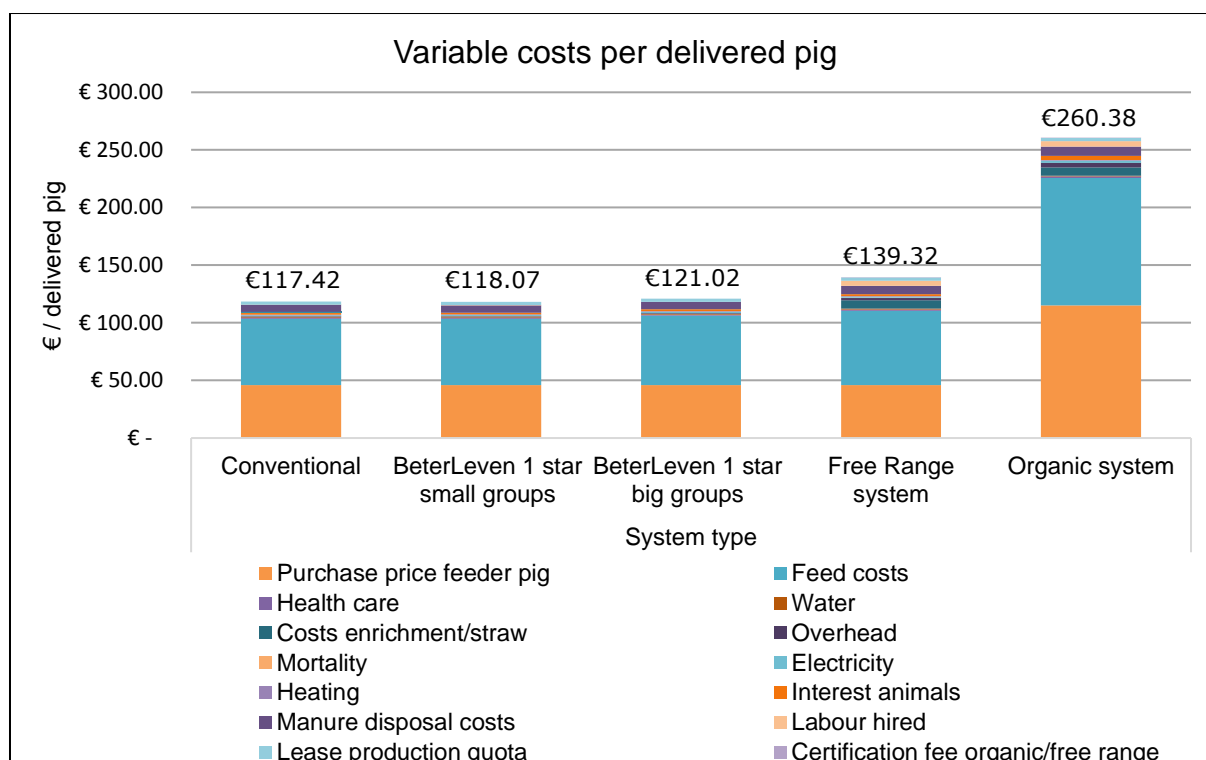


Figure 8: Variable costs (€/delivered pig) per system type

The largest share of the variable costs in all system types consists of the feed costs and the purchase of piglets. Table 10 shows the proportion of these costs compared to the total variable costs of the system. The contribution varied between the 79% and the 88% of the total variable costs of the system. The price for the purchase of piglets were equal for the first four systems, same as the feed price. The small differences in the feed costs were related to the deferring feed conversion ratio's between the systems. The organic prices for piglets and feed costs are considerably higher as the piglet and the feed are required to be certified organic and must meet strict requirements (Nahm, 2014). Nevertheless, due to the increase in other variable costs, the proportion of feed and piglet costs are comparable with the proportion of the other systems.

Table 10: Overview of the proportion of feed- and piglet costs relative to the total variable costs

	System type				
	Conventional	Beter Leven 1 star small groups	Beter Leven 1 star big groups	Free Range system	Organic system
Feed costs	57.41	57.41	60.11	64.48	110.43
Purchase price piglet	46.05	46.05	46.05	46.05	115.12
Total	103.46	103.46	106.16	110.53	225.56
<i>% of total variable costs</i>					
- Feed costs	49%	49%	50%	46%	42%
- Purchase price piglet	39%	39%	38%	33%	44%
- Total	88%	88%	88%	79%	87%

Next to the feed and piglet costs, other variables have a smaller contribution to the difference in variable costs (Figure 9). Healthcare costs are increasing moving from left to right between the systems. Between the conventional system and the Beter Leven 1 star small groups there is no difference and changing from small to big groups, there is a small increase due to mixing and more contact between a larger number of animals (Voermans, 2013). A larger increase is visible converting from the 1-star system to the free-range system and from free-range to the organic system, which is related to the assumption that systems with outdoor access have a less constant climate (Voermans, 2013). The costs for water are in the organic system slightly lower than in the other systems. The costs for enrichment in the form of straw or roughage is €0 for the conventional system, slightly higher in the 1-star system and highest in the free-range and organic system.

After the costs for feed and piglets, the costs for enrichment is the third largest cost difference. Overhead costs were equal for the first three systems and higher for the free range and organic system. The costs for mortality were slightly higher in organic as both the mortality rate and the costs per pig are higher. Electricity costs were twice as high in the organic system compared to the other systems. On the other hand, heating costs were slightly lower in the organic system as straw in combination with natural ventilation has an influence on the type of heating (Voermans, 2013). The calculated interest for animals and the manure disposal costs increased across the systems. Hired labour was considerably higher in the free-range and organic systems as more labour hours per pig are required. Lease for production quota increases slightly from conventional to organic as the amount depends on the delivered animals per year. Finally, the free-range and organic system have to pay a small certification fee.

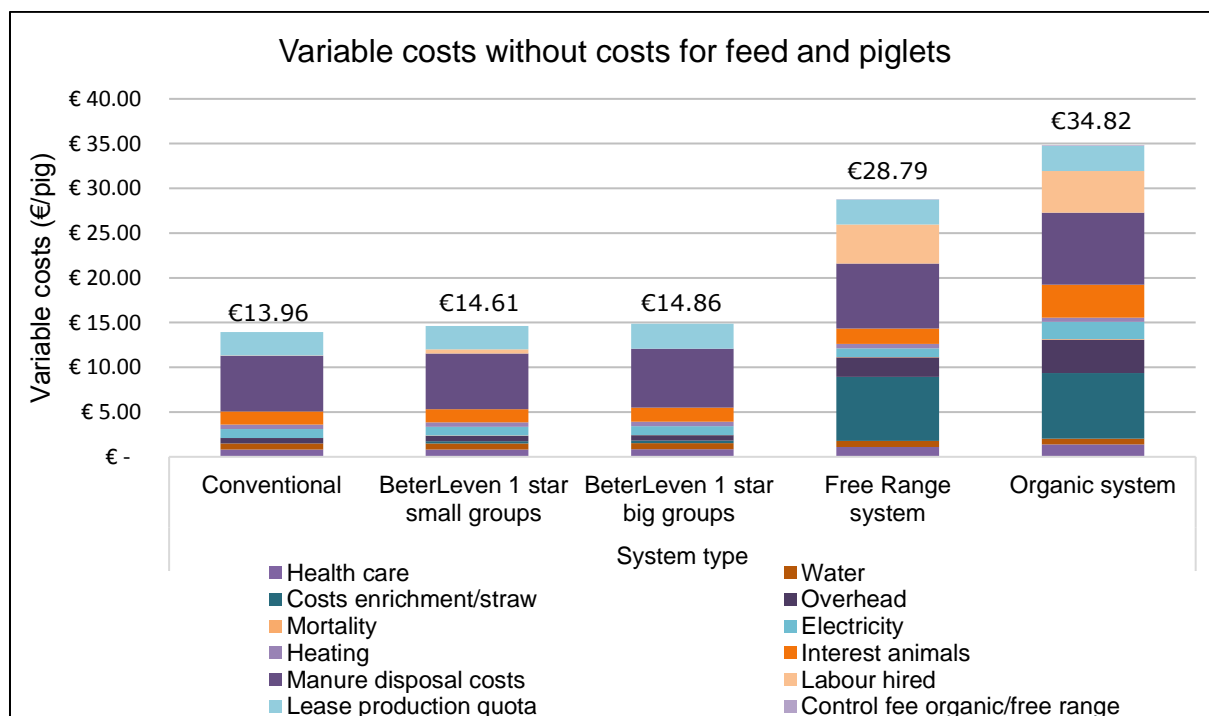


Figure 9: Variable costs (€/delivered pig) excluding feed- and piglet costs

Fixed costs are shown in Table 11 and Figure 10. Similar to the variable costs, fixed costs are highest in the organic system. However, there is no increasing trend present from the conventional to the organic system. Fixed costs are lowest in the free-range system, followed by the conventional system, Beter Leven 1 star big groups, Beter Leven 1 star small groups and finally the organic system.

Labour costs are a considerable share of the fixed cost in all systems. The highest labour costs per pig were found in the system with Beter Leven 1 star big groups, thereupon in the organic system, the free-range system and lowest in the conventional and Beter Leven 1 star small groups.

Table 11: Fixed costs (€/delivered pig) per system type

	System type				
	Conventional	Beter Leven 1 star small groups	Beter Leven 1 star big groups	Free Range system	Organic system
Labour own	5.19	5.19	5.67	5.50	5.62
Buildings, inventory, quota and land	16.41	20.18	17.30	14.99	28.59
- Depreciation buildings	3.57	4.34	3.58	2.38	5.44
- Depreciation outdoor access	0.00	0.00	0.00	2.06	2.99
- Depreciation inventory	3.43	4.27	3.83	2.91	5.20
- Depreciation NH3 emission reduction system	0.71	0.87	0.78	0.00	0.00
- Interest invested capital	6.96	8.54	7.26	6.07	11.92
- Maintenance buildings	0.89	1.09	0.90	0.59	1.36
- Maintenance inventory	0.86	1.07	0.96	0.73	1.30
- Maintenance outdoor access	0.00	0.00	0.00	0.26	0.37
Total fixed costs	21.61	25.37	22.97	20.49	34.21

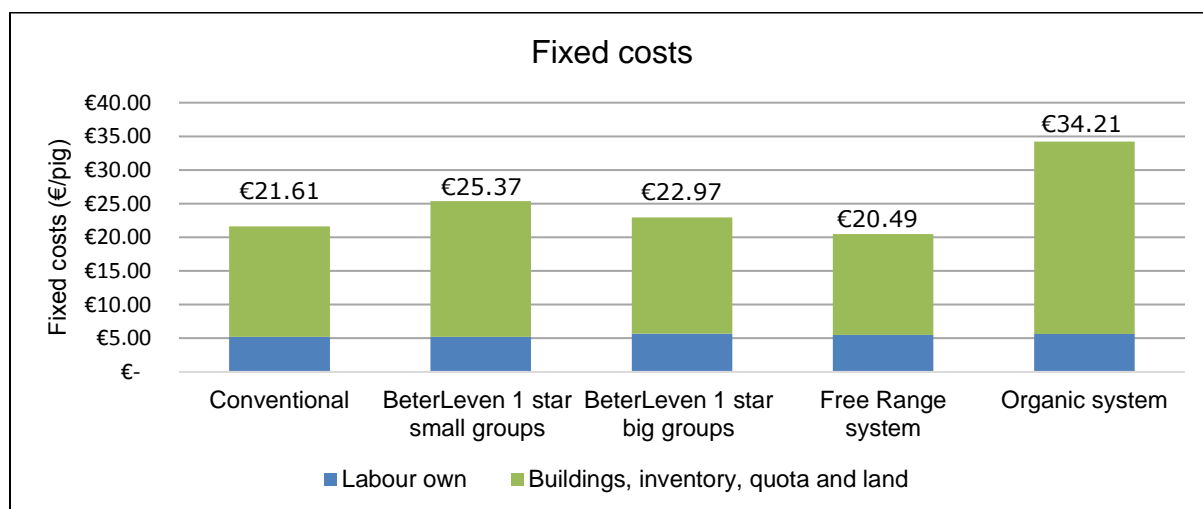


Figure 10: Fixed costs (€/delivered pig) per system type

Besides labour, fixed cost consists of buildings, inventory, quota, and land. These costs consisted of depreciation, interest and maintenance costs and are found in Table 11 and in Figure 11. The largest three contributors in all systems were interest invested capital and depreciation of inventory and buildings. They were lowest in the free-range system and highest in the organic system. Differences in buildings are affected by the air inlet system and the surface of the manure cellar and stocking density per pig (Voermans, 2013). The difference in investments in inventory are influenced by the costs of enrichment material, manure slider, type of ventilation and use of a microclimate area in the stable (Voermans, 2013). The contribution of maintenance costs for inventory and buildings was relatively low and did not considerably differ among the systems. The conventional and 1-star system had extra cost for an air scrubber which is necessary for ammonia reduction.

The free-range and organic system had additional costs for outdoor access in the form of depreciation and maintenance. These costs for outdoor access were 17,6% (€3.60) of the total fixed costs for the free-range system and 15.3% (€5.23) for the organic system. Next to the maintenance and depreciation costs, calculated interest invested capital was also influenced by the investment in outdoor access.

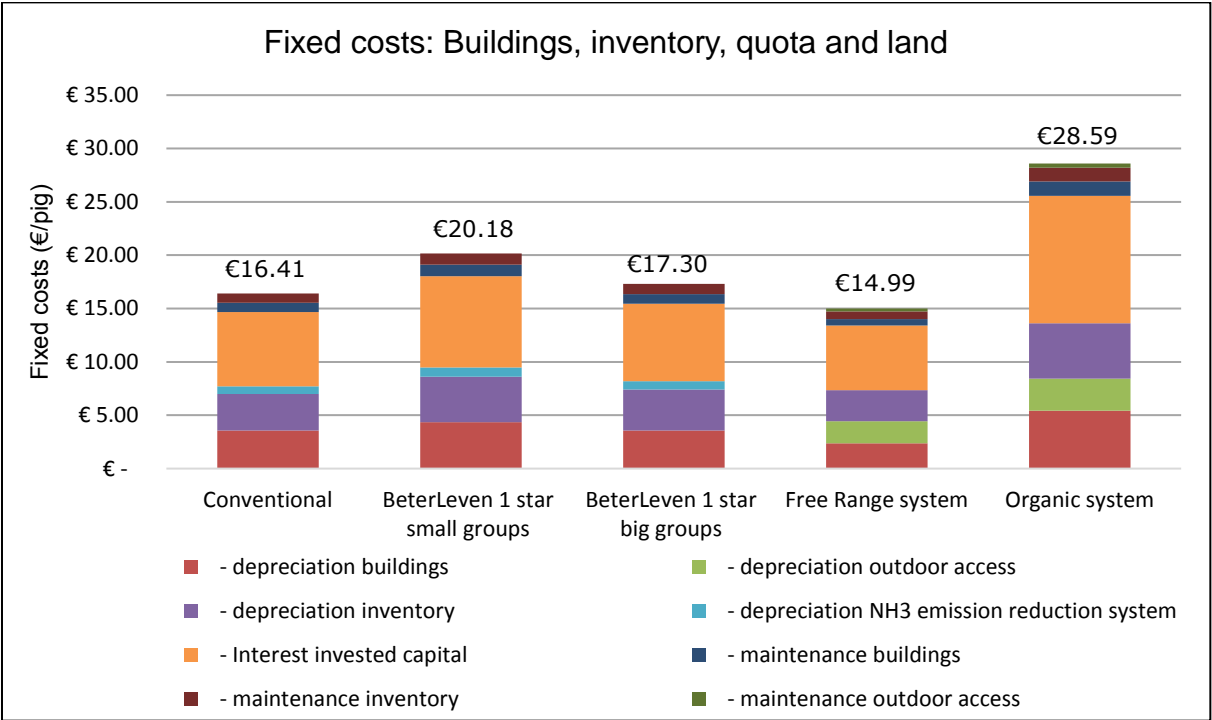


Figure 11: Overview of fixed costs (€/delivered pig) per system type excluding labour costs

The variable and fixed were combined to calculate the total production costs per delivered pig. The results are visible in Figure 12. Total costs per pig were approximately twice as high in the organic system (€294.58) compared to the other systems with total costs of €139.02 for conventional, €143.44 for 1-star small groups, €143.99 for 1-star big groups and €159.81 for free-range. For the additional costs compared to the conventional system this implies that the 1-star small groups system shows an increase of 3.2%, 1-star big groups an increase of 3.6%, free-range an increase of 15.0% and organic increased with 111.9%.

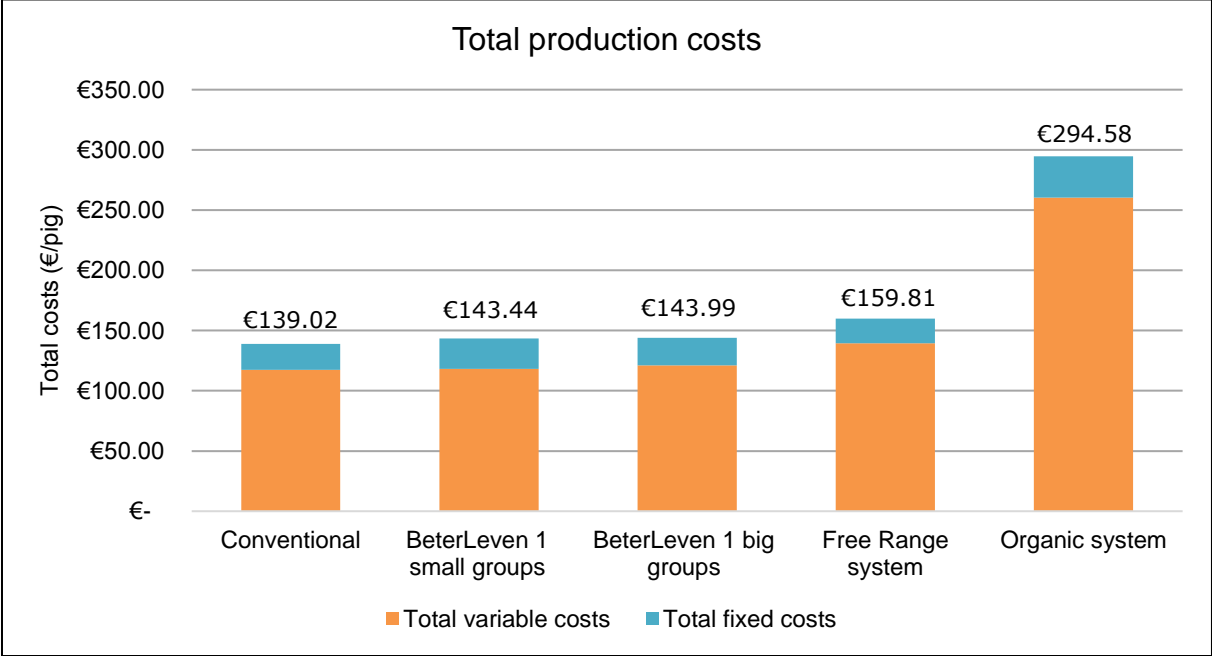


Figure 12: Total production costs (€/delivered pig) per system type

3.3 Cost-efficiency of production systems

The cost-efficiency was calculated by comparing the change in WQ Index score (ΔWQ) with the change in production costs (ΔTC). Using the conventional system as a reference, the calculated changes can be found in Table 12, the used WQ scores in Table 8, and the TC in Figure 12. In Figure 13 the WQ scores were plotted against the production costs to obtain a graphical overview of the position of the systems. It was decided to combine the small and big groups of the Beter Leven 1-star system as they had the same welfare quality score and almost equal production costs. The production costs used for 1-star in this cost-efficiency analysis is the average of the small and big group system.

Table 12: Change in WQ Index score (ΔWQ) and production costs (ΔTC in €/delivered pig), their corresponding change in % and cost-efficiency (CE) per system type

From	Beter Leven 1 star			Free range			Organic		
	ΔWQ	ΔTC	CE	ΔWQ	ΔTC	CE	ΔWQ	ΔTC	CE
Conventional	60.3 (4.8%)	4.69 (3.4%)	12.9	61.7 (4.9%)	20.79 (15%)	2.97	74.2 (5.9%)	155.56 (112%)	0.48
1-star				1.4 (0.1%)	16.10 (11%)	0.09	14.0 (1.1%)	150.87 (105%)	0.09
Free range							12.5 (0.9%)	134.77 (84%)	0.09

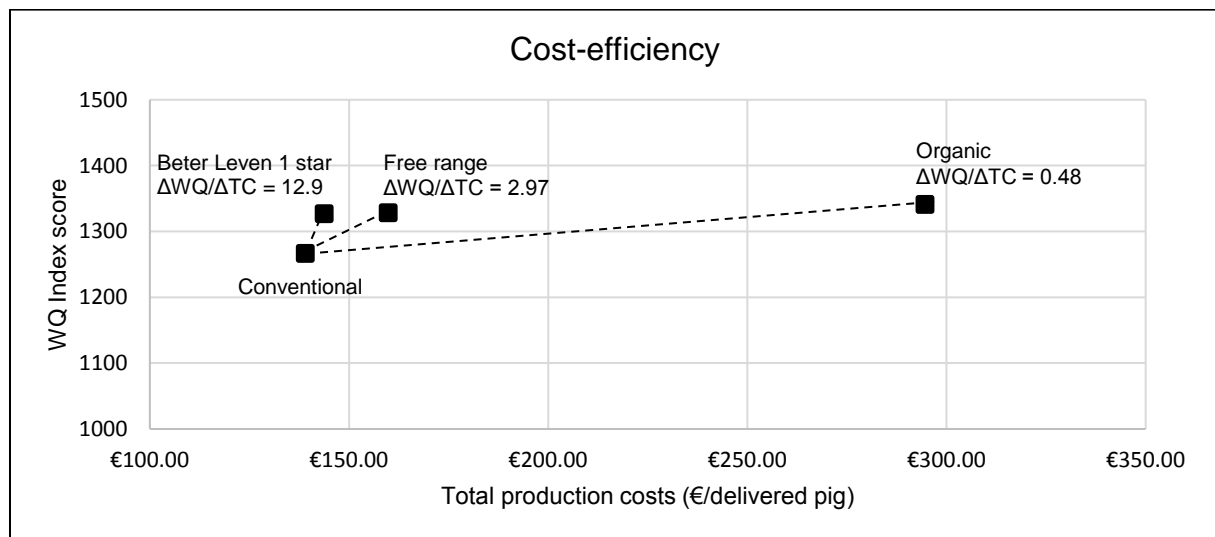


Figure 13: Plot of the total welfare quality score and the total production costs of the different system types and their corresponding cost-efficiency

Comparing the conventional and 1-star system shows an increase in WQ score of 60.3 (4.8%) and an increase in TC of €4.69 (3.4%). When converting from conventional to the free-range system, there is an increase in WQ score of 61.7 (4.9%) and a change in TC of €20.79 (15%). When organic is compared to conventional, there is a WQ score increase of 74.2 and a TC increase of €155.56. These numbers show that the increase in WQ shows a small increase up to 5.9% while the TC show a way larger increase of 112%.

Table 12 also shows the comparison to the 1-star system to the free-range and organic system. These show a ΔWQ of 1.4 (0.1%) and ΔTC of 16.10 (11%) and ΔWQ of 14 (1.1%) and ΔTC of 150.87 (105%), respectively. Comparing the free-range system and the organic system we see a change in WQ of 12.5 (0.9%) and change in TC of 134.77 (84%).

Overall, the highest cost-efficiency ($\Delta WQ/\Delta TC = 12.9$) can be observed converting from conventional to the Beter Leven 1-star system. All the other comparisons showed a considerably lower cost-efficiency. In decreasing order the cost-efficiencies were: from conventional to free-range (2.97), from conventional to organic (0.48), from free-range to organic (0.09), from 1-star to organic (0.09) and finally from 1-star to free-range (0.09). The Impact of a conversion beyond the 1-star system mostly increases costs while the WQ level almost remains the same.

3.4 Sensitivity analysis

Because some uncertainty is involved in the inputs used in this study, a sensitivity analysis was performed. This analysis is in line with the sensitivity analysis performed in Brooshooft *et al.* (2014). The sensitivity of the WQ Index score was reflected by the standard error and the maximum and minimum scores. Because no data was available for the free-range system, the standard error was estimated using the average of the 1-star and organic standard error. Maximum and minimum scores for the free-range system were estimated using the averages of the 1-star and organic system.

The sensitivity of the production costs was reflected by the best and worst case prices and their deviations from the average price. For the best and worst case prices, the highest and lowest price over the period 2013-2017 were selected from KWIN-V 2018-2019. Since data on the prices of the organic system was lacking, the difference in proportion in the conventional prices was used to calculate the organic prices. The results of these calculations are shown in Table 13.

Table 13: Results of the sensitivity analysis on WQ Index scores and production costs (€/delivered pig) per system type

	Production system			
	Conventional	1-star	Free range	Organic
WQ statistics				
WQ avg.	1272.0	1328.9	1328.2	1348.3
SE	13.9	21.0	26.5	32.0
WQ max	1471.3	1504.9	1506.9	1509.0
WQ min	1080.6	1140.1	1109.3	1078.6
Cost statistics				
Prices average	139.02	143.71	159.81	294.58
Prices worst case	161.17	166.66	183.33	345.21
Deviation mean	-22.15	-22.95	-23.52	-50.63
Prices best case	123.06	127.48	143.21	259.02
Deviation mean	15.96	16.23	16.59	35.56

The standard errors of the WQ scores and the best and worst case costs were selected to visualise the possible variation around the average scores and are shown in Figure 14. Looking to the WQ scores, it is visible that conventional is positioned underneath the other three systems regardless the scenario. The 1-star, free-range and organic system show overlap in possible WQ scores. The size of the WQ standard errors is mainly affected by the data set and only a small part because of the linkages and their weights. Changing the weights will affect all the systems in roughly the same way. The best and worst case scenario of the production costs show that the organic system is still clearly higher in the production costs than the other systems. The systems conventional, 1-star and free-range show overlap. Nevertheless, these systems use the same feed and will all show the same shift in costs when de feed and piglet prices change. Therefore, their position relative to each other won't change.

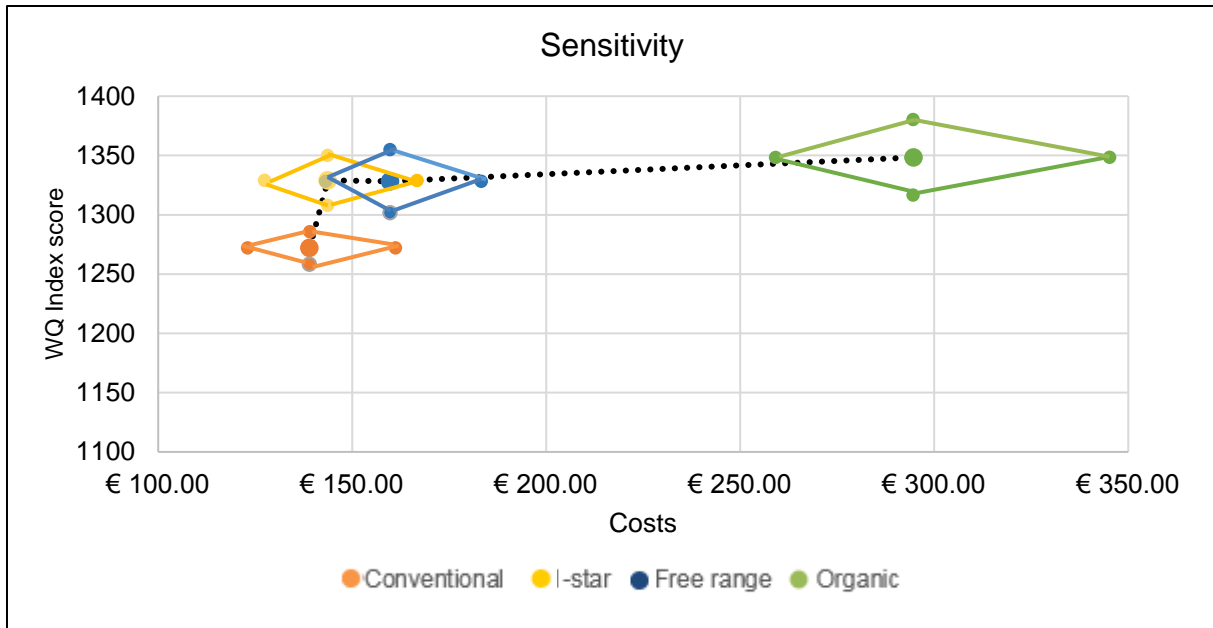


Figure 14: Results of the sensitivity analysis on WQ Index scores and production costs (€/delivered pig) per system type including the WQ standard errors (vertical) and the best- and worst case costs scenario (horizontal)

In order to compare the variable costs of the systems without the difference in feed and piglet price, a corrected version of the organic system was made. Feed and piglet prices were set equal to the other systems and this system was called 'corrected organic system'. The correction caused a change in variable costs in the costs for: purchase price feeder pig, feed costs, mortality and calculated interest animals and resulted in total variable costs of €147.48. The other variables remained unchanged. Table 14 shows the affected costs of the correction for organic prices compared to the other systems. The change in cost-efficiency is shown in Figure 15. The corrected organic system is positioned closer to the other systems relative to the organic system. The cost-efficiency increased from 0.48 to 1.72. With that, the corrected organic system still showed the lowest cost-efficiency relative to the conventional system compared with the 1-star and free-range system.

Table 14: Affected variable costs when corrected for organic feed and piglet prices in € per delivered pig of the production systems

	System type					
	Conventional	Beter Leven 1 star small groups	Beter Leven 1 star big groups	Free Range system	Organic system	Corrected Organic system
Purchase price feeder pig	46.05	46.05	46.05	46.05	115.12	46.05
Feed costs	57.41	57.41	60.11	64.48	110.43	68.52
Mortality	0.02	0.02	0.02	0.03	0.08	0.04
Calculated interest animals	1.46	1.47	1.58	1.72	3.69	1.83
Total per pig	117.42	118.07	121.02	139.32	260.38	147.48

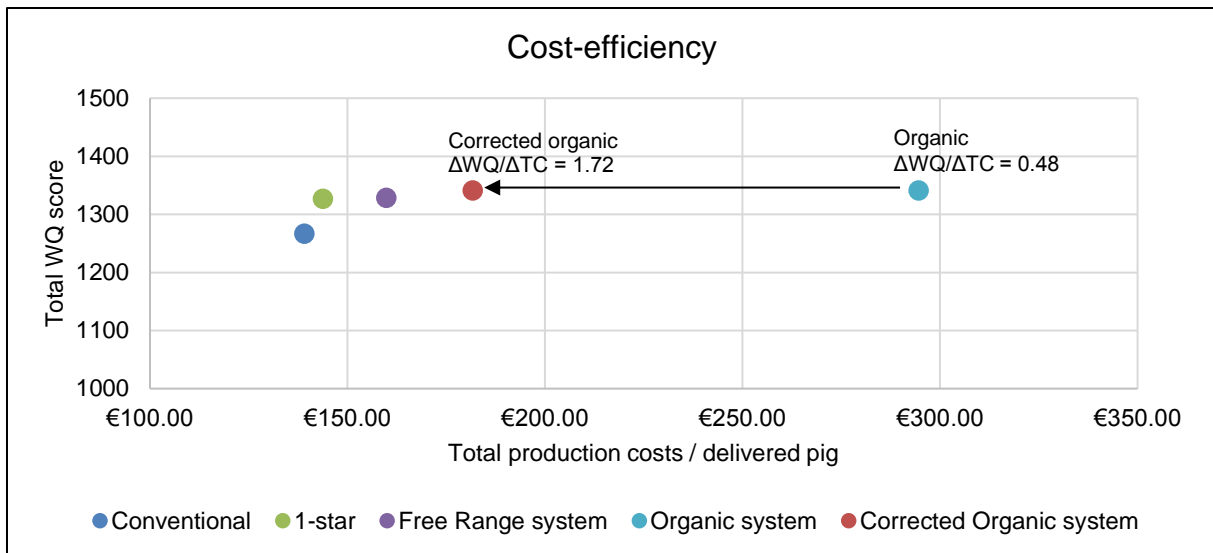


Figure 15: Plot of the total welfare quality score and the total production costs of the different system types plus the corrected organic system and their corresponding cost-efficiency

4 Discussion

The objective of this study was to gain insight in the cost-efficiency (CE) of animal welfare improvements in the Dutch growing-finishing pig sector. This CE will provide the farmers and other stakeholders with meaningful information in the decision making process for a certain pig production system type. To our knowledge this information is not provided by existing literature.

For this analysis 4 different production systems were selected: conventional (no stars), 1-star Beter Leven, free-range (2-stars) and organic (3-stars). These systems differ in their presumed AW level and range from the minimum legal standards in the Netherlands (no stars) to 3-stars Beter Leven for the organic system.

The AW level was assessed using the Welfare Quality® Assessment protocol for pigs (Welfare Quality®, 2009). In order to assess the AW level on production system level, the selected system attributes that influence AW had to be linked to the welfare measures used in the WQ protocol. The variable, fixed and total costs were calculated using the deterministic model developed by Gocsik *et al.* (2015) and the inputs were updated with the help of literature and expert opinion. The AW level and costs were combined for each system and resulted in the CE when they were compared to the conventional system.

In this section the results will be discussed, as well as the methodology and the implementations of this research. It is concluded with recommendations for further research and system development.

4.1 Results

The results of this study show that WQ index scores of these systems were very similar. The difference between the lowest and highest score is only 6.0%. The conventional system showed the lowest animal welfare level. The 1-star system had a slightly higher score, after that the free-range system and the organic system received the highest score. These last scores were almost similar and were only 1.1% higher in the organic system compared to the 1-star system.

The three system attributes that showed the highest contribution to AW were: outdoor access (23%), stocking density (19%) and bedding (18%). A perhaps surprising outcome was the negative effect of outdoor access. The organic farms scored worse on the measures linked to outdoor access. Some of these negative scores on the linked measures were expected from literature as outdoor access has a trade-off between positive and negative effects, but others differed from the expectation. Panting, lameness, wounds on body, tail biting, ruptures and hernias and investigation of the pen scored worse for the organic system while a positive effect was expected (Guy *et al.*, 2002; Spoolder, 2007; Walker and Bilkei, 2006; Vermeer, 2018).

The production costs showed increasing costs with increasing welfare and this was mostly caused by the variable costs as they explained around 85% of total costs and showed the same increasing trend. Especially in the organic system, production costs are considerably higher as the feed and piglet price are almost double. When those higher prices were corrected to the prices of the other systems, the organic system still had the highest production costs as variable costs without feed and piglet costs and fixed costs were also the highest for the organic system. Literature confirms with the fact that welfare improvements increase production costs (Bornett, Guy and Cain, 2003; Gocsik *et al.*, 2015).

Overall, the CE was highest converting from the conventional to the Beter Leven 1-star system. For the free-range and organic system, the CE was lower. In other words, the changes in system attributes in the 2- and 3-star systems result in higher production costs while the AW remains approximately the same resulting in a lower CE with the conventional system as reference. The highest CE in the 1-star system is in line with the results from Gocsik *et al.* (2016) of the broiler sector. Nevertheless, some differences could be observed regarding the AW level in this study. In this research a way larger gain in AW was observed for the 1-star system, free-range had a comparable AW level and a slightly lower AW was observed in the organic system. A possible explanation for this is the use of different breeds between the broiler systems while the pig systems use largely the same breeds. A slower growing chicken breed influences a lot of welfare measures in a positive way, resulting in a higher gain in AW level in the 1-star (Volwaard/ Puur & Eerlijk), free-range and organic systems (Gocsik *et al.*, 2016). It could also be argued that pigs in the conventional system already have a relatively high standard in AW level due to regulation and therefore are closer to the AW levels of the other systems compared to the broiler sector.

4.2 Methodology

It could be realised that the results presented in 4.1 can be influenced by the methodology used. Some of the steps contain a degree of uncertainty which might raise some questions on the credibility of these results. Therefore in this section, the linking and weighting system attributes to welfare measures, data and sensitivity and the use of the WQ protocol will be discussed in order to examine the robustness of the outcomes and hence the credibility of the conclusions.

4.2.1 Linking and weighting system attributes to welfare measures

Drawing conclusions on system level can be quite complex. As already shown in the literature research section of this research, many system attributes are affecting AW. Some animal welfare measures score worse and other best in for example in the organic system, in other words, no clear trend could be observed. Therefore, in literature, not many conclusions are drawn on system level but on measure or criteria level instead.

For the linking and weighting of the linkages between system attributes and WQ measures, the same approach was used as in Gocsik *et al.* (2016). However, the calculation of the WQ scores had some limitations because of the methodology and the data. The linking of system attributes to the WQ measures was performed through extensive literature research and checked by an expert. Therefore, it can be said that this part of the method is quite robust even though research differed in opinion on some points. The weighting of the linkages is a somewhat more arbitrary step as it was done with the help of one expert. It is possible that other experts would divide the weights slightly different which results in some uncertainty regarding the contribution of different system attributes. In the sensitivity analysis on the WQ scores of this study, the standard error of the calculated was taken to analyse the uncertainty and did not contain a more thorough analysis on the adjustment of the weights of the linkages due to time constraints. In the research of Gocsik *et al.* (2016) such a sensitivity analysis was performed and changing the weights did not lead to a difference in the ranking of the systems. Although no sensitivity analysis was performed on this part, it is expected that the ranking in this study is also quite robust as the effect would be roughly the same in all system types. If for example stocking density was founds more important, in all systems the score for stocking density would increase. For the systems with a lower stocking density the increase will be slightly higher but the overall ranking will remain the same. Therefore it could be stated that the establishment of the linkages is more important than the weighting. However, a possible change in ranking in this study after adjustment of weights cannot be excluded. Hence, a separate sensitivity analysis should be done on the weights of the linkages is advised.

Another aspect that made the linking and weighting process more complex is the existence of interactions between the system attributes. For example systems with outdoor access have at the same time a lower stocking density and therefore various measures are linked to multiple attributes. This also ensures that a certain system attribute, like tail docking, does not have equal scores for systems that both perform tail docking which indicates that some of these interactions are indirectly included. Nevertheless, literature about these interactions is very scarce and could therefore not be included as a separate variable.

The welfare quality protocol normally uses the WQ measure scores to calculate criteria and principle scores. In this research, however, this was not an option as not all measures could be linked to a system attribute. The different effects of the system attributes could be best visualised with the measure scores as more and various trends could be observed. For example, the criteria absence of disease consist out of various measures and not all measures display the same trend across the systems. Only using the measure scores has a downside, namely that the weighting of the measures into criteria scores is not present in this study. In this research, all measures contribute evenly to the WQ Index scores while in WQ not every measure would receive the same weight. Therefore, criteria that consist of more measures could be over-represented and influence the outcome more than they would in WQ. Also, WQ considers the measures to be independent of each other while clearly some interactions exist. For example, pleurisy and pneumonia share many risk factors and therefore they tend to occur together (Stärk, Pfeiffer and Morris, 1998; Sanchez-Vazquez, 2013).

4.2.2 Data and sensitivity

Some uncertainty in the results could be assigned to the data used for the calculation of the welfare scores. Data was already collected before 2012 and the Beter Leven quality mark set foot in the pig sector in 2014 (Peet *et al.*, 2018). Therefore farms had to be assigned to the star ranking with the help of extra data. Results could differ if the data was more recent and farms truly were in possession of a certain label of Beter Leven and really produce according to those standards instead of accidentally meet the standards. It is expected that conditions in the relative young systems at the time, like organic, are nowadays much improved. Besides that, sample sizes for the production systems were not equal and there was no access to data for the free-range system what makes the results less certain.

A sensitivity analysis was performed to analyse the impact of uncertainty regarding the estimated CE's, particularly the robustness of the relative differences. The standard errors of the WQ scores were mainly influenced by the scores from the dataset as there is no difference in the linked welfare scores and their assigned weight. These scores of the 1-star, free-range and organic system were almost similar and the sensitivity analysis showed overlap in the possible WQ score. For the production costs, sensitivity was visualised with best and worst case price scenario's. The sensitivity analysis shows some overlap however, in practice the overlap is not present as the worst case price scenario would apply to all systems. This makes the ranking of the overall results quite robust.

4.2.3 Assessment of animal welfare using WQ

In this research, the Welfare Quality protocol was used to assess the level of animal welfare in the production systems. The Welfare Quality® protocol is a project stimulated by the European Commission in order to develop a tool that measures the level of animal welfare systematically and is currently a commonly used tool. The measures of the protocol were mostly based on animal-based measures what has the advantage that the outcome of the effects of housing design and management and their interaction are all measured. Nevertheless, the protocol resulted in some counter-intuitive outcomes in this study. The results of this study imply that the level of AW in the free range and organic system is not distinctly higher than the 1-star system while these systems claim to be animal friendlier. Hence, the question is legitimate if WQ covers all welfare aspects.

For example, with some of the welfare attributes like bedding and outdoor access, a trade-off of positive and negative effects exist. Measuring the positive effects of AW improvements seems rather difficult and might be under-represented in the protocol (Vermeer, 2018). Only the measures exploratory behaviour, QBA and to a lesser extent social behaviour (score based on negative social behaviour) show the positive effects of AW improvements. The motivation of the pigs for certain attributes could be checked with operate conditioning (Dawkins, 1988). In the data set used in this study, WQ scores show that outdoor access has a negative effect on AW. Nonetheless, pigs still choose to go outside while they could stay inside. This shows that pigs get a certain satisfaction from being outside that is not measurable with WQ. The same goes for straw bedding. The beneficial effects of straw are described in multiple articles and pigs will lay on straw when they are given the choice (Tuytens, 2005). However, it is not a direct measure in the protocol and has to show its indirect effect through multiple measures, still resulting in a presumably underestimated effect. This example shows that certain positive aspects might be underrepresented in the protocol.

To elaborate on the possible shortcomings of WQ the theory of the Maslow pyramid of needs can be used (Maslow, 1943). This theory is a motivational theory in psychology that depicts different levels of human needs, often depicted as hierarchal levels within a pyramid shape (Figure 16) (McLeod, 2007). This pyramid consists out of basic needs, psychological needs and self-fulfilment needs. Pigs are highly intelligent animals and therefore it could be debated if they should also be met in the levels beyond the basic needs as they share multiple cognitive ability's with humans (Mendl, Held and Byrne, 2010). In the WQ protocol the main focus is on the basic needs of the animal and barely on the psychologically needs. Most of the basic needs can be visible through animal-based measures. The psychological needs on the other hand, can be measured more difficult, especially when the protocol has to remain non-invasive and inexpensive.

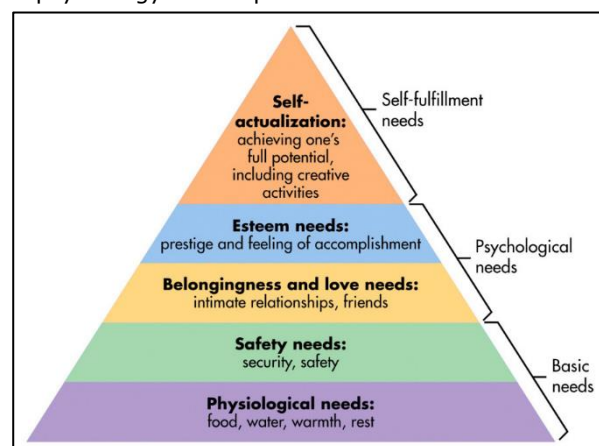


Figure 16: Maslow pyramid of needs

Because these psychological needs are not clearly present in the protocol it could give a distorted picture of the actual welfare level. The expectation is that the top segment systems respond more to the psychological needs to the pig via better enrichment material, outdoor access, and straw bedding. If this is the case, the current WQ protocol could underestimate the AW level of the top segment systems.

In the end, the question arises if the increased production costs accompanying the higher star systems are worth it regarding AW. More research on the psychological wellbeing of pigs is needed in order to answer this question.

4.2.4 Conclusion

To conclude, the methodology of this research contains some uncertainties. Nevertheless, it is expected that the overall ranking of the systems regarding their CE is quite robust. Hence, The relative ranking of the production systems is far more important than the absolute values. However, some counterintuitive outcomes were found in this study. The results could imply that the claimed level of AW in the top segment systems are in reality lower. Another explanation would be the possible underestimation of the AW level in the higher star systems as those systems seem better in meeting the psychological needs and WQ does not fully include those beyond basic needs. This would imply the WQ protocol has some shortcomings and improvements are required on the field of the measurement of psychological well-being of the pig.

4.3 Implications

One of the objectives of this study is to offer the farmers and other stakeholders an extra information source when it comes to production system choice. Converting to the 1-star system comes with the highest cost-efficiency, also interpretable as best value for the farmers investment. Converting from conventional to the 1-star system would be a good option as it improves the AW level at relative low costs, namely €4,42 per pig. Nevertheless, for a farmer with 3000 pig places this comes down to approximately €42.000 a year which is a considerable amount of extra costs. Farmers only adopt systems with a higher AW level if they are compensated for the extra costs and they receive a price premium for their product.

The perception of the consumers plays an important role when it comes to price premiums. Consumers need to be willing to pay this price premium for an animal friendlier product. For example, outdoor access is perceived by consumers as an important point for AW (Vanhonacker *et al.*, 2008). This research shows that with the current WQ protocol, outside access has a negative influence on AW. In the research of Gocsik *et al.* (2016) the level of AW even slightly decreased for the organic system compared to extensive outdoor. Hence, some contradictions exist when it comes to the perception of AW of consumers and the AW assessments in scientific analyses. Chances are high that this consumer perception also influenced the development of the Beter Leven quality mark and the ranking system as for example outside access is incorporated in the requirements of 2 and 3 stars (Beter Leven Keurmerk, 2019b).

Even though the highest cost efficiency was found for the 1-star system, the possibly low willingness-to-pay for that product could form an obstacle. Consumers' willingness-to-pay is therefore crucial in the improvement of AW in pig production systems.

The Dutch retail can play an important role in this problem. The past years retail began to remove the conventional meat products. At first, the products were replaced by products in between the conventional and 1-star products like the "*Kip van morgen*" concept in the broiler sector (Gocsik *et al.*, 2016). After the chicken products, the pig products followed with the "*Varken van morgen*" concept. Currently, more and more retailers are only selling products with at least 1-star. By removing the conventional choice, the 1-star product becomes the cheapest and most sold option. The *Kip van morgen* increased prices between 15-20% compared to the conventional option but did not cause a decrease in chicken meat sales. The problem of the low willingness to pay for the 1-star products, is now avoided through policy decisions. It shows that such policy decisions of company's or the government is needed to address the welfare issues.

4.4 Future research and system development

From this study some issues for further research can be derived. To begin with, AW assessment methods should be developed that are able to make conclusions on production system level. Furthermore, more research is needed for mapping the interactions between the WQ measures and the system attributes. In this way, the welfare levels of the production systems could be estimated more accurately. WQ could improve their protocol by focussing on improved measures that estimate the psychological level of AW as it seems underrepresented in the current method. In addition, the development of future AW assessment methods should emphasize the estimation of the AW on the psychological level.

The use of more recent data, equal group sizes and adding data of free-range farms could lead to more accurate estimations of the AW level. As already mentioned in 4.2.1 a separate sensitivity analysis should be performed to analyse the influence of weight changes for the established linkages in order to estimate the robustness of the welfare ranking. Besides that, other welfare assessment tools than WQ could be used to compare the differences with the current results. Although it is not expected that the ranking of the systems would be different due to the inability of existing tools in measuring the psychological needs and the extensiveness of the WQ protocol. Moreover, these results of the Dutch sector could be compared to future research of the cost-efficiency in other countries. This research was performed in the growing-finishing pig sector. Future research could be focussing on the sow and piglet sector and could show different results as they are exposed to different environmental conditions.

More extensive research is necessary on consumer perception and their willingness to pay in relation to the extra production costs. The willingness to pay of consumers for certain system attributes could be compared to the AW level measured by scientific assessment methods. At the moment there are some contradicting views on AW comparing consumer and scientific perception. To tackle the contradicting views of consumers the focus should lay with consumer education on the welfare problems in the pig sector in order to create a public opinion that is more in line with current research on AW. With more extensive research on the willingness to pay of consumers for certain AW aspects an estimation of the possible revenues of the farmer and retail could be made. With that information the cost-efficiency of this research could be transformed to a benefit-cost analysis. This will give the farmer and other stakeholders more insight into the viability and margins of these systems in relation to the AW gain. The willingness of consumers is important but will not always lead to a shift in purchase behaviour. Retail has therefore a crucial role in improving AW. By making the collective decision to only sell pig meat with 1-star or beyond, comparable with broiler products in 2016, a collective problem could be addressed.

Future system development should focus on the aspects that improve AW that have the highest cost-efficiency like stocking density and bedding. Also experimenting with different breeds based on their resilience and aggression could be beneficial to the AW level. Breeding not only based on productivity but also on AW could lead to fewer welfare problems due to health problems or lesions especially in the free-range and organic sector.

5 Conclusion

Results of this research show that WQ index scores of these systems were very similar. The difference between the lowest and highest score is only 6%. The organic system scores best with a total WQ score of 1348 after that in decreasing order the free-range system (1336), the 1-star system (1329) and finally the conventional system (1272). The three system attributes that showed the highest contribution were: outdoor access, stocking density and bedding. Different trends were observed between the systems for the measure scores and the system attribute scores. The variable costs showed an increasing trend from conventional to the organic system with similar amounts for the conventional, 1-star small groups and 1-star big groups system, a small increase in the free-range system and twice as high for the organic system. Fixed costs are lowest in the free-range system, followed by the conventional system, Beter Leven 1-star big groups, Beter Leven 1-star small groups and finally the organic system. The variable and fixed costs combined, resulted in the total costs to be approximately twice as high in the organic system (€294.58) compared to the other systems with total costs of €139.02 for conventional, €143.44 for 1-star small groups, €143.99 for 1-star big groups and €159.81 for free-range. Overall, the highest cost-efficiency ($\Delta WQ/\Delta Costs = 12.9$) can be observed converting from conventional to the Beter Leven 1-star system. All the other comparisons showed a considerably lower cost-efficiency: from conventional to free-range (2.97) and from conventional to organic (0.48).

The sensitivity analysis showed that the relative ranking of costs was quite robust. Likewise, when corrected for the higher prices for feed and piglets in the organic system, the ranking did not change even though the differences were smaller. The welfare scores of 1-star, free-range, and organic system overlapped and there was no clear ranking. However, it was already stated that the scores were practically the same.

Concluding, it can be stated that the highest cost-efficiency can be obtained converting from the conventional system to the 1-star system. Beyond the 1-star system the CE decreased due to a similar welfare level and reasonable increased costs. However, due to presumed limitations of the WQ protocol it is possible that welfare scores of the organic and free-range system were underestimated.

Acknowledgements

This master thesis was part of my MSc. program Animal Sciences at Wageningen University. In my chosen specialisation global and sustainable production, economics is an important subject and a subject of interest for me personally. In order to expand my knowledge on the area of economics, I chose to do a professional track within my master. The professional track Business and management consisted out of following a set of courses and writing an extra minor thesis, in my case, at the Business Economics group.

The subject of this thesis resembles the combination of economics with global and sustainable production and was therefore directly in my area of interest. My thanks go to my supervisors Helmut Saatkamp and Herman Vermeer for their expertise, guidance, and feedback during the past months. Mainly because of them it has been a pleasurable and informative period. Besides them, I would like to thank Robert Hoste and Luuk Vissers for their help and knowledge.

References

- Agrimatie (2018) *Prijzen in varkensvleesketen lager dan een jaar eerder*. Available at: <https://www.agrimatie.nl/themaResultaat.aspx?subpubID=2232§orID=2420&themaID=3596&indicatorID=2414> (Accessed: 25 February 2019).
- AHAW (EFSA Panel on Animal Health and Welfare). (2014). Scientific Opinion concerning a Multifactorial approach on the use of animal and non-animal-based measures to assess the welfare of pigs. *EFSA Journal*, 12(5), 3702.
- Alban, L., Petersen, J. and Busch, M. (2015) 'A comparison between lesions found during meat inspection of finishing pigs raised under organic/free-range conditions and conventional, indoor conditions', *Porcine Health Management*, 1(1), p. 4.
- Andersen, I. L. Nævdal, E., Bakken, M., & Bøe, K. E. (2004) 'Aggression and group size in domesticated pigs, *Sus scrofa*: "when the winner takes it all and the loser is standing small"', *Animal Behaviour*, 68(4), pp. 965–975.
- Appleby, M. C. (2004) 'What Should We Do about Animal Welfare?', *The Quarterly Review of Biology*, 76(1), pp. 56–56.
- ASR real estate (2019) *Grondprijzen*. Available at: <https://asrrealestate.nl/landelijk-vastgoed/grondprijzen> (Accessed: 25 February 2019).
- Avé, M. Bosman, T., Rousou, X., Stadig, L., Van Winden, R., Van Der Woude, L., Klingen, K., Smid, A., Schouls, J. (2012) *The effects of outdoor and indoor housing on pig health*. Available at: <https://www.boerengroep.nl/wp-content/uploads/2011/12/Effects-on-Pig-Health-ACT-Student-Report-for-Boerengroep.pdf> (Accessed: 19 February 2019).
- Averós, X. Brossard, L., Dourmad, J. Y., de Greef, K. H., Edge, H. L., Edwards, S. A., and Meunier-Salaün, M. C. (2010) A meta-analysis of the combined effect of housing and environmental enrichment characteristics on the behaviour and performance of pigs, *Applied Animal Behaviour Science*, 127(3–4), pp. 73–85.
- Backus, G. B. C. (2004) Castratie : via verdoofd naar volledig stoppen. *V-focus*, 6(5), 38-39.
- Backus, G. B. c. and Schans, J. W. Van Der (2000) *Pig farmers in a dialogue with society*. (No. 6.00. 06). LEI.
- Beattie, V. E. Sneddon, I. A., Walker, N., and Weatherup, R. N. (2001) 'Environmental enrichment of intensive pig housing using spent mushroom compost', *Animal Science*, 72(01), pp. 35–42.
- Beattie, V. E., Walker, N. and Sneddon, I. A. (1996) 'An investigation of the effect of environmental enrichment and space allowance on the behaviour and production of growing pigs', *Applied Animal Behaviour Science*, 48(3–4), pp. 151–158.
- Beter Leven Keurmerk (2019a) *Beter Leven dieren met het Beter Leven Keurmerk - Beter Leven Keurmerk*. Available at: <https://beterleven.dierenbescherming.nl/> (Accessed: 4 February 2019).
- Beter Leven Keurmerk (2019b) *Factsheet Varkens*. Available at: https://beterleven.dierenbescherming.nl/fileupload/pdf_sterren_consumenten/Tekst_BLK_varkens.pdf (Accessed: 6 February 2019).
- von Borell, E. and Sørensen, J. T. (2004) 'Organic livestock production in Europe: aims, rules and trends with special emphasis on animal health and welfare', *Livestock Production Science*, 90(1), pp. 3–9.
- Bornett, H. L. I., Guy, J. H. and Cain, P. J. (2003) 'Impact of Animal Welfare on Costs and Viability of Pig Production in the UK', *Journal of Agricultural and Environmental Ethics*, 16(2), pp. 163–186.
- Broom, D. M. (1991) 'Assessing welfare and suffering', *Behavioural Processes*, 25(2–3), pp. 117–123.
- Broom, D. M. and Fraser, A. F. (2015) 'The welfare of pigs.', in *Domestic animal behaviour and welfare*, pp. 272–280.
- Brooshoof, S., De Jong, I. C., Gocsik, E., and Saatkamp, H. W. (2014). Economic optimization of animal welfare in the Dutch broiler sector using the Welfare Quality® assessment protocol for poultry. In Book of Abstracts Workshop Meeting of the Welfare Quality Network (pp. 10-10). Clinical Vet Science, University of
- Bristol UK. Bussink, F. (2018) *Prijzen voor hooi en stro gaat omhoog*. Available at: <https://www.boerderij.nl/Rundveehouderij/Achtergrond/2018/3/Prijzen-voor-hooi-en-stro-gaat-omhoog-262677E/> (Accessed: 25 February 2019).
- Cleveland-Nielsen, A., Nielsen, E. . and Ersbøll, A. (2002) 'Chronic pleuritis in Danish slaughter pig herds', *Preventive Veterinary Medicine*, 55(2), pp. 121–135.
- Cornale, P., Macchi, E., Miretti, S., Renna, M., Lussiana, C., Perona, G. and Mimosi, A. (2015) 'Effects of stocking density and environmental enrichment on behavior and fecal corticosteroid levels of pigs under commercial farm conditions', *Journal of Veterinary Behavior*, 10(6), pp. 569–576.
- Cornale, P., Macchi, E., Miretti, S., Renna, M., Lussiana, C., Perona, G., Battaglini, L. M. (2015) *Proceedings of the 49th Congress of the International Society for Applied Ethology*. Edited by T. Yasue et al. The Netherlands: Wageningen Academic Publishers.

- Courboulay, V. and Delarue, E. (2009) 'Welfare assessment in 82 pig farms : effect of animal age and floor type on', *Animal Welfare*, 18(4), pp. 515–521.
- Dawkins, M. S. (1988) 'Behavioural deprivation: A central problem in animal welfare', *Applied Animal Behaviour Science*, 20(3–4), pp. 209–225.
- Ernst, K. Ekkelboom, M., Kerksen, N., Smeets, S., Sun, Y., Yin, X. (2018) *Play behavior and environmental enrichment in pigs*. Available at: [https://www.wur.nl/upload_mm/f/0/8/fcc4320a-6ec8-4df0-9661-c0f13287a63e_Kristi Ernst e.a. %282018%29 Play behavior and environmental enrichment in pigs. Concretizing principle-based regulation.pfd.pdf](https://www.wur.nl/upload_mm/f/0/8/fcc4320a-6ec8-4df0-9661-c0f13287a63e_Kristi_Ernst_e.a._%282018%29_Play_behavior_and_environmental_enrichment_in_pigs_Concretizing_principle-based_regulation.pfd.pdf) (Accessed: 20 February 2019).
- Etterlin, P. E. Ytrehus, B., Lundeheim, N., Heldmer, E., Österberg, J., and Ekman, S. (2014) 'Effects of free-range and confined housing on joint health in a herd of fattening pigs', *BMC Veterinary Research*, 10(1), p. 208.
- European Union (1998) *Council Directive 98/58/EC* of 20 July 1998 concerning the protection of animals kept for farming purposes. Official Journal L, 221(08), 08.
- Ewbank, R. and Bryant, M. J. (1972) 'Aggressive behaviour amongst groups of domesticated pigs kept at various stocking rates', *Animal Behaviour*, 20(1), pp. 21–28.
- Fredriksen, B. Lium, B. M., Marka, C. H., Mosveen, B., and Nafstad, O. (2008) 'Entire male pigs in farrow-to-finish pens—Effects on animal welfare', *Applied Animal Behaviour Science*, 110(3–4), pp. 258–268.
- Fu, L. Li, H., Liang, T., Zhou, B., Chu, Q., Schinckel, A. P., and Huang, R. (2016) 'Stocking density affects welfare indicators of growing pigs of different group sizes after regrouping', *Applied Animal Behaviour Science*, 174, pp. 42–50.
- Gillman, C. E. KilBride, A. L., Ossent, P., & Green, L. E. (2008) 'A cross-sectional study of the prevalence and associated risk factors for bursitis in weaner, grower and finisher pigs from 93 commercial farms in England', *Preventive Veterinary Medicine*, 83(3–4), pp. 308–322.
- Gocsik, É., Saatkamp, H. W., De Lauwere, C. C., and Lansink, A. O. (2014) 'A Conceptual Approach for a Quantitative Economic Analysis of Farmers' Decision-Making Regarding Animal Welfare', *Journal of Agricultural and Environmental Ethics*. Springer Netherlands, 27(2), pp. 287–308.
- Gocsik, É. Lansink, A. O., Voermans, G., and Saatkamp, H. W. (2015) 'Economic feasibility of animal welfare improvements in Dutch intensive livestock production: A comparison between broiler, laying hen, and fattening pig sectors', *Livestock Science*, 182, pp. 38–53.
- Gocsik, É., Brooshooft, S. D., de Jong, I. C., and Saatkamp, H. W. (2016) 'Cost-efficiency of animal welfare in broiler production systems: A pilot study using the Welfare Quality® assessment protocol', *Agricultural Systems*, 146, pp. 55–69.
- Guy, J. H., Rowlinson, P., Chadwick, J.P. and Ellis, M. (2002) 'Health conditions of two genotypes of growing-finishing pig in three different housing systems: implications for welfare', *Livestock Production Science*, 75(3), pp. 233–243.
- Hogg, A. and Switzer, W. P. (1978) *Atrophic Rhinitis History*. Available at: https://research.libraries.wsu.edu:8443/xmlui/bitstream/handle/2376/12655/em4175_1978.pdf?sequence=1 (Accessed: 14 February 2019).
- INRA (2018) *Welfare Quality® scoring system*. Available at: <http://www1.clermont.inra.fr/wq/index.php?id=protocol&prod=> (Accessed: 21 February 2019).
- Jensen, T. (2009) 'Causes of and predisposing risk factors for leg disorders in growing-finishing pigs.', *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 4(010).
- Jensen, T. Nielsen, C.K., Vinther, J. and D'Eath, R.B. (2012) 'The effect of space allowance for finishing pigs on productivity and pen hygiene', *Livestock Science*, 149(1–2), pp. 33–40.
- Juska, R., Juskiene, V. and Leikus, R. (2013) 'The influence of a free-range housing system on pig growth, carcass composition and meat quality', *Journal of Applied Animal Research*, 41(1), pp. 39–47.
- KWIN (2018) *KWIN Kwantitatieve Informatie Veehouderij 2018-2019*. 29th edn. Wageningen: Wageningen UR Livestock Research.
- Lassen, J., Sandøe, P. and Forkman, B. (2006) 'Happy pigs are dirty! – conflicting perspectives on animal welfare', *Livestock Science*, 103(3), pp. 221–230.
- Lensink, B. J. Ofner-Schröck, E., Ventorp, M., Zappavigna, P., Flaba, J., Georg, H. and Bizeray-Filoché, D. (2013) 'Lying and walking surfaces for cattle, pigs and poultry and their impact on health, behaviour and performance', in *Livestock housing*. The Netherlands: Wageningen Academic Publishers, pp. 75–92.
- Maes, D. Verdonck, M., Deluyker, H., and de Kruif, A. (1996) 'Enzootic pneumonia in pigs', *Veterinary Quarterly*, 18(3), pp. 104–109.
- Maes, D. Deluyker, H., Verdonck, M., Castryck, F., Miry, C., Vrijens, B., and De Kruif, A. (2000) 'Herd factors associated with the seroprevalences of four major respiratory pathogens in slaughter pigs from farrow-to-finish pig herds', *Veterinary Research*, 31(3), pp. 313–327.

- Maes, D. G. Maes, D. G., Deluyker, H., Verdonck, M., de Kruif, A., Ducatelle, R., Castryck, F., and Vrijens, B. (2001) 'Non-infectious factors associated with macroscopic and microscopic lung lesions in slaughter pigs from farrow-to-finish herds.', *The Veterinary record*. British Medical Journal Publishing Group, 148(2), pp. 41–6.
- Maslow, A. H. (1943) 'A theory of human motivation.', *Psychological Review*, 50(4), pp. 370–396.
- McLeod, S. (2007) 'HCC Certificate in Counselling Skills Maslow's Hierarchy of Needs', (December), p. 3.
- Mendl, M., Held, S. and Byrne, R. W. (2010) 'Pig cognition', *Current Biology*, 20(18), pp. R796–R798.
- Meyer-Hamme, S. E. K., Lambertz, C. and Gauly, M. (2016) 'Does group size have an impact on welfare indicators in fattening pigs?', *animal*, 10(01), pp. 142–149.
- Miao, Z. H., Glatz, P. C. and Ru, Y. J. (2004) 'Review of Production, Husbandry and Sustainability of Free-range Pig Production Systems', *Asian-Australasian Journal of Animal Sciences*, 17(11), pp. 1615–1634.
- Michelsen, J. Hamm, U., Wynen, E., & Roth, E. (1999) 'The European Market for Organic Products: Growth and Development'. Universität Hohenheim - Stuttgart Hohenheim. Available at: <http://orgprints.org/8486/> (Accessed: 4 February 2019).
- Moinard, C. Moinard, C., Mendl, M., Nicol, C. J., & Green, L. E. (2003) 'A case control study of on-farm risk factors for tail biting in pigs', *Applied Animal Behaviour Science*, 81(4), pp. 333–355.
- Nahm, K. H. (2014) 'Nutrition and Feeding of Organic Poultry', *Asian-Australasian Journal of Animal Sciences*, 22(2), pp. 305–306.
- Nannoni, E., Valsami, T., Sardi, L., & Martelli, G. (2014) 'Tail Docking in Pigs: A Review on its Short- And Long-Term Consequences and Effectiveness in Preventing Tail Biting', *Italian Journal of Animal Science*, 13(1), p. 3095.
- Park, H.-S., Min, B. and Oh, S.-H. (2017) 'Research trends in outdoor pig production - A review.', *Asian-Australasian journal of animal sciences*, 30(9), pp. 1207–1214.
- Pearce, G. P. (1999) 'Epidemiology of enteric disease in grower-finisher pigs: a postal survey of pig producers in England.', *The Veterinary record*, 144(13), pp. 338–342.
- Peet, G. F. V. van der, van der Meer, R. W., van Leeuwen, H. D., and van Wageningen-Lucardi, S. R. M. (2018) 'Monitoring integraal duurzame stallen : Peildatum 1 januari 2018'. Available at: <https://library.wur.nl/WebQuery/wurpubs/542061> (Accessed: 4 February 2019).
- Petersen, H. (2008) 'Prevalence of clinical signs of disease in Danish finisher pigs', *Veterinary Record*, 162(12), pp. 377–382.
- Petersen, V., Simonsen, H. B. and Lawson, L. G. (1995) 'The effect of environmental stimulation on the development of behaviour in pigs', *Applied Animal Behaviour Science*, 45(3–4), pp. 215–224.
- POV (2017) *Factsheet varkenshouderij*. Available at: http://pov.nl/fileadmin/user_upload/Factsheet_varkenshouderij_november_2017.pdf (Accessed: 25 February 2019).
- Quiniou, N., Courboulay, V., Salaün, Y., & Chevillon, P. (2010) Impact of the non castration of male pigs on growth performance and behaviour-comparison with barrows and gilts. In Conference at the 61st Annual Meeting of the European Association for Animal Production, Heraklion, Crete Island, Greece, 7pp.
- Rodenburg, T. B. and Koene, P. (2007) 'The impact of group size on damaging behaviours, aggression, fear and stress in farm animals', *Applied Animal Behaviour Science*, 103(3–4), pp. 205–214.
- Roepstorff, A. and Nansen, P. (1998) *FAO Animal Health Manual: Epidemiology, diagnosis and control of helminth parasites of swine, (Vol. 3)*. Rome: FAO
- Ruis, M., Pinxterhuis, I. and Vrolijk, M. (2010) 'Update welzijnsprestaties biologische veehouderij', (No. 317). Wageningen UR Livestock Research.
- Sanchez-Vazquez, M. J. (2013) *Epidemiological Investigations Utilizing Industry Abattoir Data - a Study in Finishing Pigs*, (Doctoral dissertation, Utrecht University).
- Schild, S. L. A., Rousing, T., Jensen, H. E., Barington, K., and Herskin, M. S. (2015) 'Do umbilical outpouchings affect the behaviour or clinical condition of pigs during 6 h housing in a pre-transport pick-up facility?', *Research in Veterinary Science*. W.B. Saunders, 101, pp. 126–131.
- Schmolke, S. A., Li, Y. Z. and Gonyou, H. W. (2003) 'Effect of group size on performance of growing-finishing pigs', *Journal of Animal Science*, 81(4), pp. 874–878.
- Schröder-Petersen, D. . and Simonsen, H. . (2001) 'Tail Biting in Pigs', *The Veterinary Journal*, 162(3), pp. 196–210.
- Scott, K., Chennells, D. J., Campbell, F. M., Hunt, B., Armstrong, D., Taylor, L., and Edwards, S. A. (2006) 'The welfare of finishing pigs in two contrasting housing systems: Fully-slatted versus straw-bedded accommodation', *Livestock Science*, 103(1–2), pp. 104–115.

- Da Silva, C. A., Manteca, X. and Dias, C. P. (2016) 'Needs and challenges of using enrichment materials in the pig industry', *Semina: Ciências Agrárias*, 37(1), pp. 525–536.
- Silva, M., Exadaktylos, V., Ferrari, S., Guarino, M., Aerts, J. M., and Berckmans, D. (2009) 'The influence of respiratory disease on the energy envelope dynamics of pig cough sounds', *Computers and Electronics in Agriculture*. Elsevier, 69(1), pp. 80–85.
- Simonsen, H. B. (1995) 'Effect of Early Rearing Environment and Tail Docking on Later Behaviour and Production in Fattening Pigs', *Acta Agriculturae Scandinavica, Section A - Animal Science*. Taylor & Francis Group, 45(2), pp. 139–144.
- Smith, W. (1993) 'A study of adventitious bursitis of the hock of the pig'. (Doctoral dissertation, University of Edinburgh).
- Sørensen, J. T., Edwards, S., Noordhuizen, J., and Gunnarsson, S. (2006) Animal production systems in the industrialised world, *Rev Sci Tech*, 25(2), 493-503.
- Spoolder, H. A. (2007) 'Animal welfare in organic farming systems', *Journal of the Science of Food and Agriculture*. John Wiley & Sons, Ltd, 87(15), pp. 2741–2746.
- Spoolder, H. A. M., Edwards, S. A. and Corning, S. (1999) 'Effects of group size and feeder space allowance on welfare in finishing pigs', *Animal Science*. Cambridge University Press, 69(03), pp. 481–489.
- Stärk, K. D. C., Pfeiffer, D. U. and Morris, R. S. (1998) 'Risk factors for respiratory diseases in New Zealand pig herds', *New Zealand Veterinary Journal*. Taylor & Francis Group, 46(1), pp. 3–10.
- Street, B. R. and Gonyou, H. W. (2008) 'Effects of housing finishing pigs in two group sizes and at two floor space allocations on production, health, behavior, and physiological variables¹', *Journal of Animal Science*. Oxford University Press, 86(4), pp. 982–991.
- Sundrum, A. and Weißmann, F. (2005) *Organic pig production in free range systems. Special Issue 281 Bundesforschungsanstalt für Landwirtschaft, library.wur.nl*. Available at: <https://library.wur.nl/WebQuery/wurpubs/fulltext/44699#page=49> (Accessed: 19 February 2019).
- Temple, D., Courboulay, V., Manteca, X., Velarde, A., and Dalmau, A. (2012) 'The welfare of growing pigs in five different production systems: assessment of feeding and housing', *animal*. Cambridge University Press, 6(04), pp. 656–667.
- Turner, S. P., Ewen, M., Rooke, J. A., and Edwards, S. A. (2000) 'The effect of space allowance on performance, aggression and immune competence of growing pigs housed on straw deep-litter at different group sizes', *Livestock Production Science*. 66(1), pp. 47–55.
- Turner, S. P., Horgan, G. W. and Edwards, S. A. (2001) 'Effect of social group size on aggressive behaviour between unacquainted domestic pigs', *Applied Animal Behaviour Science*. 74(3), pp. 203–215.
- Tuytens, F. A. M. (2005) 'The importance of straw for pig and cattle welfare: A review', *Applied Animal Behaviour Science*. 92(3), pp. 261–282.
- Vanhonacker, F., Verbeke, W., Van Poucke, E., & Tuytens, F. A. (2008) 'Do citizens and farmers interpret the concept of farm animal welfare differently?', *Livestock Science*. 116(1–3), pp. 126–136.
- Veissier, I., Butterworth, A., Bock, B., & Roe, E. (2008) 'European approaches to ensure good animal welfare', *Applied Animal Behaviour Science*. 113(4), pp. 279–297.
- Vermeer, H. M., de Greef, K. H. and Houwers, H. W. J. (2014) 'Space allowance and pen size affect welfare indicators and performance of growing pigs under Comfort Class conditions', *Livestock Science*. 159, pp. 79–86.
- Vermeer, H., van Reenen, C. and Spoolder, H. (2012) *Vereenvoudiging Welfare Quality® protocol voor varkens*. (No. 622). Wageningen UR Livestock Research.
- Voermans, G. (2013) *A comparison of the mid-term economic feasibility of animal welfare systems on pig finishing farms in the Netherlands*. Available at: <http://edepot.wur.nl/273451> (Accessed: 4 March 2019).
- Walker, P. K. and Bilkei, G. (2006) 'Tail-biting in outdoor pig production', *The Veterinary Journal*, 171(2), pp. 367–369.
- Webster, A. J. (2001) 'Farm animal welfare: the five freedoms and the free market'. *The veterinary journal*, 161(3), 229-237.
- Weerd, H. A. Van de . Docking, C. M. Day, J. E. Breuer, K. and Edwards, S. A. (2006). (2006) 'Effects of species-relevant environmental enrichment on the behaviour and productivity of finishing pigs', *Applied Animal Behaviour Science*, 99(3–4), pp. 230–247.
- van de Weerd, H. A. and Day, J. E. L. (2009) 'A review of environmental enrichment for pigs housed in intensive housing systems', *Applied Animal Behaviour Science*. 116(1), pp. 1–20.
- Welfare Quality® (2009) *Welfare Quality® assessment protocol for pigs*. *Welfare Quality Consortium, Lelystad, Netherlands*. Available at: http://www.welfarequalitynetwork.net/media/1018/pig_protocol.pdf (Accessed: 4 February 2019).

Weng, R. ., Edwards, S. . and English, P. . (1998) 'Behaviour, social interactions and lesion scores of group-housed sows in relation to floor space allowance', *Applied Animal Behaviour Science*, 59(4), pp. 307–316.

Wet Dieren (2015) *Wet Dieren 2015*. Available at: <https://wetten.overheid.nl/BWBR0030250/2015-02-01#Hoofdstuk2> (Accessed: 22 March 2019).

Widowski, T. (2010) 'The Physical Environment and Its Effect on Welfare', in. Springer, pp. 137–164.

Quinn, A. J. (2014) 'Limb health in pigs: The prevalence and risk factors for lameness, limb lesions and claw lesions in pigs, and the influence of gilt nutrition on indicators of limb health', *Livestock Science*, 156(1–3), pp. 53–61.

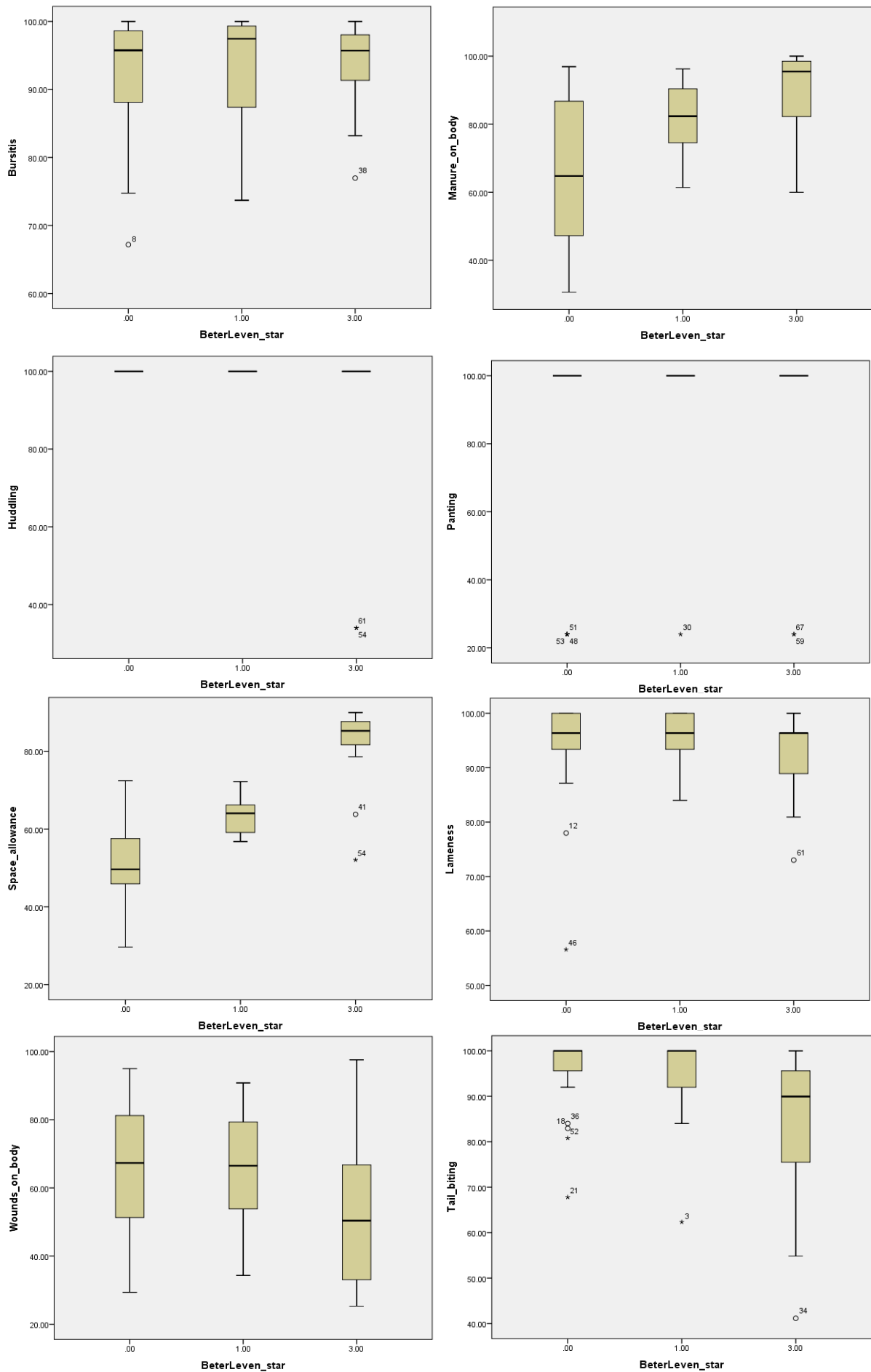
Wolter, B F. Ellis, M. Curtis, S E. Augspurger, N R. Hamilton, D N. Parr, E N. Webel, D M (2001) 'Effect of group size on pig performance in a wean-to-finish production system.', *Journal of Animal Science*, Press, 79(5), p. 1067.

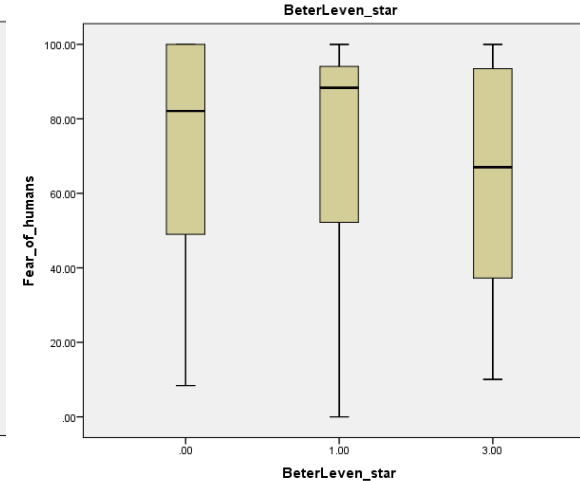
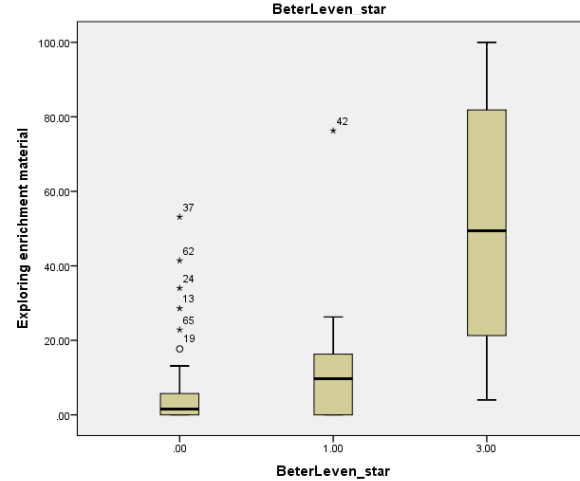
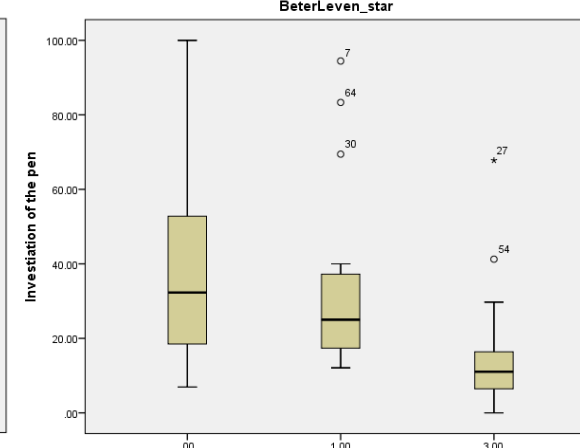
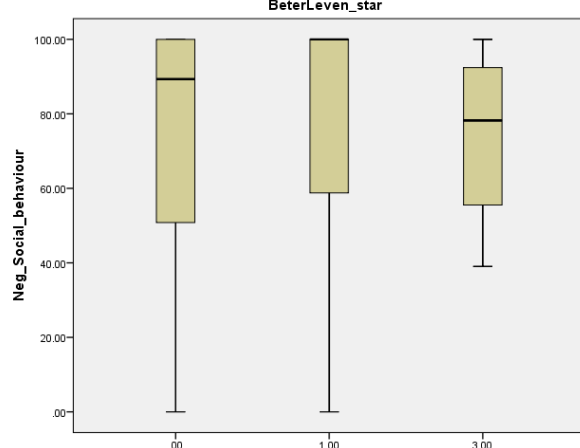
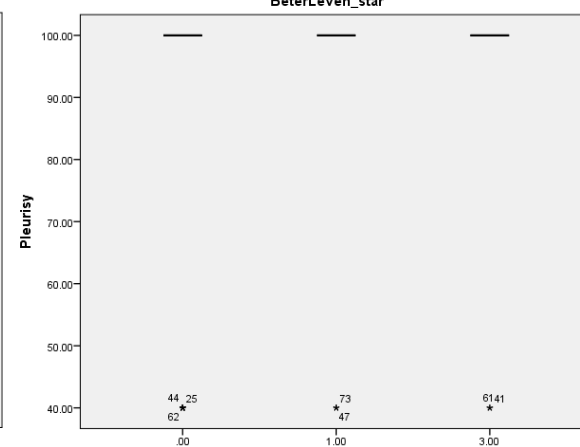
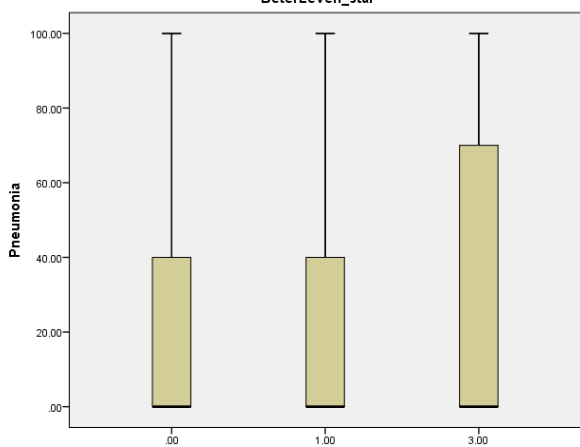
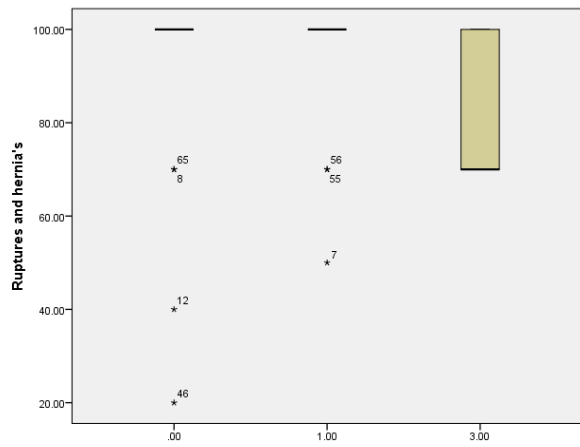
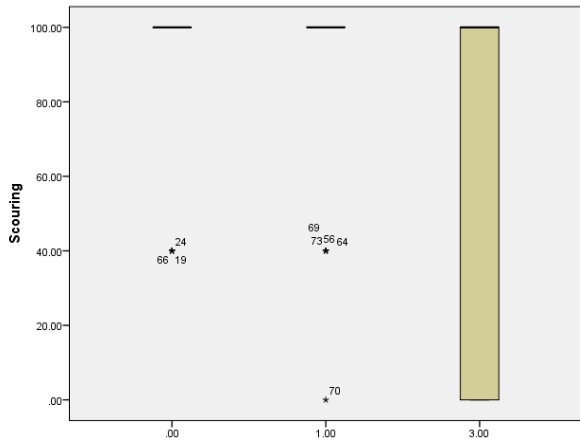
Personal contacts:

Hoste, Robert, Economist pig production Wageningen Economic Research (2018)

Vermeer, Herman, DLO researcher Wageningen Livestock Research (2018)

Appendix I. Boxplots of selected welfare measure scores per system types





Appendix II. Formulas used in economic model from (Gocsik *et al.*, 2015)

Technical

Average present fattening pigs (APFP) = Occupation rate * Stable capacity

Length growth period (days) = Weight mutation start-finish (kg) / Grow/animal/day (gram) *1000

Used feed per delivered animal (kg) = Weight mutation start-finish (kg) * Feed conversion (kg)

Circulation rounds a year = 365/ (Length growth period (days) + Length cleaning/drying period (days))

Delivered animals per year = Average present fattening pigs (APFP) * Circulation rounds a year

Total labour required (working hours) = Average present fattening pigs (APFP) * Labour required (hrs/1.000 fattening pigs/year) / 1000

Hired labour required (working hours) = Total labour required (working hours) - Total working hours farmer available

Total indoor stable surface (m²) = Stable capacity * surface/animal place

Total outdoor stable surface (m²) = surface/animal place outdoor * Stable capacity

Variable costs

Netto purchase price feed (per 100 kg) = Standard feed price (per 100 kg) + Cluster discount (per 100 kg) + Yearly quantity discount (per 100 kg)

Feed costs (€/delivered pig) = Netto purchase price feed (per 100 kg) / Used feed per delivered animal (kg / 100)

Purchase price feeder pig = Piglet price (€/23 kg animal)

Costs enrichment/straw (€/delivered pig) = Used roughage/straw/sawdust (gram per animal per day) * Length growth period (days) / 1000 * Price wood fiber/straw/roughage per kg

Average value pig = Purchase price feeder pig + (Feed costs + Health care costs + Water costs + Costs enrichment/straw + Overhead) / 2

Mortality (€/delivered pig) = Average value pig * Mortality rate

Interest animals = Average value pig * Calculated interest livestock * Length growth period (days) / 365

Manure disposal costs (€/delivered pig) = Produced manure dry feed/APFP (ton) * Manure disposal per ton * Average present fattening pigs (APFP) / Delivered animals per year

Labour costs hired (€/delivered pig) = Hired labour required (working hours) * Labour costs hired per hour / Delivered animals per year

Lease production quota (€/delivered pig) = Average present fattening pigs (APFP) * 8 / Delivered animals per year

Control fee organic/free range (€/delivered pig) = fee organic/free range / Delivered animals per year

Fixed cost

Labour costs (€/delivered pig) = Total working hours farmer available * Labour costs farmer per hour / Delivered animals per year

Depreciation buildings = Depreciation investments in buildings / Delivered animals per year

Depreciation outdoor access = Depreciation investment in outdoor access / Delivered animals per year

Depreciation inventory = Depreciation investments in inventory / Delivered animals per year

Depreciation NH₃ emission reduction system = Depreciation investment in air scrubber / Delivered animals per year

Interest invested capital = (Total book value begin year * Calculated interest Invested capital (%)) / Delivered animals per year

Maintenance buildings = Replacement value investments in buildings * Replacement value buildings (%) * Maintenance buildings (%) / Delivered animals per year

Maintenance inventory = Replacement value investments in inventory * Replacements value inventory (%) * maintenance inventory (%) / Delivered animals per year

Maintenance outdoor access = Replacement value investment in outdoor access * Replacement value outdoor access (%) * Maintenance outdoor access (%) / Delivered animals per year