

Eliciting Agents' Behaviour using Scenario-Based Questionnaire in Agent-Based Dairy Supply Chain Simulation

Journal:	<i>Journal of Simulation</i>
Manuscript ID	TJSM-2019-OP-0181.R1
Manuscript Type:	Original Paper
Date Submitted by the Author:	03-Feb-2020
Complete List of Authors:	Utomo, Dhanan; Heriot Watt University, Centre for Sustainable Road Freight Onggo, Bhakti; University of Southampton Faculty of Business and Law, Southampton Business School Eldridge, Stephen; Lancaster University Management School, Management Science Daud, Andre; Universitas Padjadjaran, Faculty of Animal Husbandry Tejaningsih, Safitri; Universitas Padjadjaran, Faculty of Animal Husbandry
Keywords:	agent-based simulation, data collection, decision rule elicitation, supply chain, agriculture
Abstract:	A scenario-based questionnaire is a survey method that aims to identify the respondents' decision rules using their responses to a series of scenarios. It is rarely used in agent-based modelling and simulation (ABMS) with most researchers preferring a survey with closed questions as the data collection method. This is particularly true for ABMS studies in agri-food supply chains. In our paper, we design a scenario-based questionnaire to elicit the behaviour of agents in ABMS and apply it in a dairy supply chain case. Our findings suggest that respondents respond well to a scenario-based questionnaire as it relates more closely to their actual decision-making process. Furthermore, our experiment shows that the decision rules extracted using a scenario-based questionnaire improve ABMS validity.

1
2
3 **Eliciting Agents' Behaviour using Scenario-Based Questionnaire in**
4 **Agent-Based Dairy Supply Chain Simulation**
5
6
7

8 D.S. Utomo^a, B.S.S. Onggo^{b*}, S. Eldridge^c, A.R. Daud^d and S.
9 Tejaningsih^d
10
11
12

13 *^aCentre for Sustainable Road Freight, Heriot Watt University, Edinburgh, UK; ^bCentre*
14 *for Operational Research, Management Sciences and Information Systems, University*
15 *of Southampton, Southampton, UK; ^cDepartment of Management Science, Lancaster*
16 *University, Lancaster, UK; ^dFaculty of Animal Husbandry, Padjadjaran University,*
17 *Jatinangor, Indonesia*
18
19
20
21
22

23 *corresponding author: B.S.S. Onggo, 02/4007 University Road, University of
24 Southampton, Southampton SO17 1BJ, UK, b.s.s.onggo@soton.ac.uk
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Eliciting Agents' Behaviour using Scenario-Based Questionnaire in Agent-Based Dairy Supply Chain Simulation

A scenario-based questionnaire is a survey method that aims to identify the respondents' decision rules using their responses to a series of scenarios. It is rarely used in agent-based modelling and simulation (ABMS) with most researchers preferring a survey with closed questions as the data collection method. This is particularly true for ABMS studies in agri-food supply chains. In our paper, we design a scenario-based questionnaire to elicit the behaviour of agents in ABMS and apply it in a dairy supply chain case. Our findings suggest that respondents respond well to a scenario-based questionnaire as it relates more closely to their actual decision-making process. Furthermore, our experiment shows that the decision rules extracted using a scenario-based questionnaire improve ABMS validity.

Keywords: agent-based modelling and simulation; data-collection; decision rule elicitation; supply chain; dairy

1. Introduction

Agent-based modelling and simulation (ABMS) is an Operational Research (OR) method that has gained popularity as a decision support tool owing to its ability to relate individual behaviours to the emerging patterns of the behaviour of the system as a whole. In common with other simulation paradigms, ABMS aims to understand a problem entity. A problem entity can be something realistic (e.g., a real system or phenomenon, an ongoing policy) or something that is not happening currently (e.g., a proposed system, idea, or a planned policy) (Sargent, 2013). Depending on the type of problem entity being studied, Gilbert (2004) classifies ABMS into *realistic models*, which are those that aim to incorporate realistic mechanisms only, and *artificial models*, which are those that aim to also incorporate unreal mechanisms. The type of mechanisms modelled will, consequently, influence the ABMS development and validation techniques used. Moss (2008) classifies ABMS development and validation

1
2
3 techniques into theory-driven (“theoretical”) and evidence-driven (“empirical”). By
4
5 empirical, we mean that the ABMS development and validation process takes into
6
7 account real data (Schutte, 2010). Most artificial models are developed theoretically
8
9 while a realistic ABMS can be developed theoretically, empirically, or via a
10
11 combination of the two. The focus of the study in this paper is on realistic models.
12
13

14 The objective of our paper is to examine the usefulness of the *scenario-based*
15
16 *questionnaire*, an alternative type of questionnaire survey method, in eliciting the
17
18 behaviours of agents in an ABMS model. In contrast with the prevalence of the use of
19
20 standard questionnaire survey (*i.e.*, survey using close-ended questionnaire) in ABMS
21
22 studies, the use of the scenario-based questionnaire is very limited. Nevertheless, in
23
24 investigating respondents’ beliefs, attitudes, or judgments a scenario-based
25
26 questionnaire is believed to have high internal and external validity (Atzmüller and
27
28 Steiner, 2010), and therefore we can expect that it can produce agents’ behaviours that
29
30 resemble those of the real actors. Hence, our paper contributes to the methodological
31
32 research in eliciting the behaviours of agents in an ABMS model.
33
34
35
36

37 In order to demonstrate how the scenario-based questionnaire can be deployed,
38
39 we used the example of an ABMS of a dairy supply chain in West Java, Indonesia.
40
41

42 Initially, we developed a baseline ABMS that was realistic using a combination
43
44 of the theoretical and empirical approaches. It used the findings of previous literature
45
46 (*i.e.*, theoretical) supplemented by domain expert interviews and parameterisation of the
47
48 results of a standard questionnaire survey (*i.e.*, empirical). The aim of the ABMS was to
49
50 predict the cattle and cow population and the daily milk production in the case study
51
52 site. We then validated the baseline ABMS using concept of operational validity
53
54 (Sargent, 2013). In parallel with this process, we derived our scenario-based
55
56 questionnaire from the agents’ decision rules of the baseline ABMS. Primary data
57
58
59
60

1
2
3 collection was then carried out using our scenario-based questionnaire and the survey
4
5 results were used to revise the agents' decision rules in our ABMS. Finally, we
6
7 validated the revised models to enable a comparison between their validity and that of
8
9 the baseline model.
10

11
12 The baseline ABMS was developed based on the theories in the previous
13
14 literature and was parameterised using a standard questionnaire survey as commonly
15
16 used in agricultural supply chain quantitative studies, which limits its contribution
17
18 beyond that of a standard ABMS case study. However, our study provides novelty and
19
20 a greater contribution in that: (i) we propose a process to design a scenario-based
21
22 questionnaire from a baseline ABMS; (ii) we quantitatively show that the additional
23
24 steps we propose are beneficial in producing a model with higher operational validity
25
26 (more predictive); and (iii) we identify and discuss the benefits of using a scenario-
27
28 based questionnaire to complement a standard questionnaire survey when developing an
29
30 ABMS. To our knowledge, the use of a scenario-based questionnaire in an ABMS study
31
32 is very limited, and guidelines for designing and deploying it in an ABMS study are
33
34 absent.
35
36
37
38
39

40 The remainder of this paper is organised as follows. In section 2, we present the
41
42 literature review to position our research in the context of behaviour elicitation method
43
44 in ABMS. In section 3, we describe the dairy supply chain case that we have selected
45
46 for our research. In section 4, we describe the sequence followed in order to develop a
47
48 scenario-based questionnaire and the survey process. We then describe, in section 5, the
49
50 data analysis and the derivation of the elicited decision rules. In section 6, we discuss
51
52 the insights obtained from the scenario-based questionnaire survey and the simulation
53
54 experiment process to test the effects of these elicited decision rules on the model
55
56 validity. Finally, we end our paper with a conclusion.
57
58
59
60

2. Literature Review

This section discusses the role of standard questionnaire survey in simulation studies, especially ABMS. We also discuss the drawbacks of a standard questionnaire survey and the potential of a scenario based questionnaire as a complementary approach.

2.1. Data collection in simulation studies

The collection and use of real-world data for ABMS development and validation is still one of the challenges facing future ABMS research (Banks, 2002; Bruch and Atwell, 2015; Hahn, 2013; Heath, Hill, & Ciarallo, 2009) and simulation modelling in general (Barlas, Heavey, & Dagkakis, 2015; Onggo and Hill, 2014; Perera and Liyanage, 2000), with up to 40% of the research time being time spent on data collection (Onggo, Hill, & Brooks, 2013; Perera and Liyanage, 2000; Trybula, 1994). There is a clear need for structured data collection methodologies in simulation (Skoogh and Johansson, 2008) and numerous data collection methodologies have been proposed to address this need (e.g. Onggo and Hill (2014), Skoogh and Johansson (2008) and Perera and Liyanage (2000)) and Barlas, et al. (2015) provide a useful review of their use. Models with higher level of detail exhibit a higher probability to cause data collection issues (Perera and Liyanage, 2000, Robinson, 1994, p. 68). This suggests that collecting and using data for ABMS models is arguably are more challenging still.

Depending on the requirements of the model, data in simulation modelling can be grouped into three types: contextual data to understand the problem situation; data for model realisation; and data for model validation (S. Robinson, 2004). In ABMS, understanding a problem situation includes an understanding as to how agents process information and make decisions.

1
2
3 There are many data collection methods in simulation. In their review, Onggo
4 and Hill (2014) identify data collection methods that have been used in simulation,
5 including role play games, questionnaires, censuses, and historical data. Similarly, a
6 variety of data collection methods have also been applied in the context of ABMS,
7 including surveys with a close-ended questionnaire, case studies, stylized facts,
8 participant observation, role-playing games, field and laboratory experiments,
9 interviews and expert knowledge (An, 2012; Janssen and Ostrom, 2006; D. T. Robinson
10 et al., 2007; Smajgl, Brown, Valbuena, & Huigen, 2011; Utomo, Onggo, & Eldridge,
11 2018; Yang and Gilbert, 2008). However, Utomo, et al. (2018) identify the standard
12 questionnaire survey method is very common, especially in the field of agri-food supply
13 chains, which encompasses the example used in this paper.
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

30 *2.2. The use of questionnaire survey data collection in simulation studies*

31 A standard questionnaire survey provides a quantitative method for collecting data on
32 individuals using a series of questions, each with pre-defined sets of possible answers
33 (closed-ended questions) (D. T. Robinson, et al., 2007; Smajgl, et al., 2011). When only
34 a fraction of the population is sampled then it is called as sample survey (D. T.
35 Robinson, et al., 2007). When the responses are collected from the entire population,
36 then it is called as a census (Smajgl, et al., 2011).
37
38
39
40
41
42
43
44
45

46 The standard questionnaire survey is one of the important research methods used
47 in ABMS (Janssen and Ostrom, 2006). A standard questionnaire survey's responses can
48 be used to determine coefficients and constraints in an equation-based ABMS based on
49 microeconomic theory (D. T. Robinson, et al., 2007). For example, Happe, Hutchings,
50 Dalgaard, & Kellerman (2011) use a farm survey to identify available resources and
51 their potential and then create a linear optimisation matrix that describes plant and
52 livestock production activities. Responses from a standard questionnaire survey are also
53
54
55
56
57
58
59
60

1
2
3 useful in generating statistical descriptions of the agents' attributes in a population
4
5 (Smajgl, et al., 2011). For example, Morgan, Brown, & Daigneault (2015) use
6
7 responses from a standard questionnaire survey to estimate the key characteristics of
8
9 demographics, income, risk tolerance and current farm practices of human actors.
10
11
12 Another use of the data from a standard survey is to construct a typology of the agents.
13
14 For example, Valbuena, Verburg, & Bregt (2008) use data concerning demographics,
15
16 perceptions and farm structures from their survey to classify clusters of agents. These
17
18 examples illustrate that the data obtained via a standard questionnaire survey are mainly
19
20 used for model realisation purposes. Their use in understanding the problem situation
21
22 and validation, particularly to elicit and validate agents' decision rules, is very limited.
23
24
25

26
27 However, this outcome is not without reason and a standard questionnaire
28
29 survey is indeed valuable. Researchers usually design a standard questionnaire survey
30
31 based on previous theories and this can effectively maintain the correspondence
32
33 between existing theories and the research results (Eldabi, Irani, Paul, & Love, 2002).
34
35 This is confirmed for example by Utomo, et al. (2018) who show that, in the field of
36
37 agriculture supply-chain, most ABMS that employs a standard questionnaire survey also
38
39 make some references to a specific theory. This is clearly important because if the
40
41 ABMS have no relationship with the previous theories at all, then its validity can be
42
43 considered to be low (Jager and Janssen, 2002).
44
45
46

47
48 However, these advantages are not without weaknesses. Some authors suggest
49
50 that sometimes a standard questionnaire survey can focus very much on theory
51
52 verification and post-decision rationalisation (namely, testing whether a theory can
53
54 explain a decision that has been made by real actors) (Eldabi, et al., 2002). This focus is
55
56 also one of the reasons why standard questionnaire surveys predominantly use a
57
58 retrospective self-report format and close-ended questions. If a process of designing a
59
60

1
2
3 standard questionnaire places too much emphasis on theory verification then,
4
5 sometimes, the concepts described in the questionnaire can be meaningless for the
6
7 respondents (i.e., they never consider these concepts when making decisions) (Yang and
8
9 Gilbert, 2008). At the same time, the close-ended questions format does not provide
10
11 much opportunity for the respondents to express their point of view.
12
13

14
15 Regarding decision rules elicitation, some authors such as Janssen and Ostrom
16
17 (2006) consider that responses to a standard questionnaire are prone to memory loss
18
19 bias (i.e., they are only reliable for very salient events). Others, such as D. T. Robinson,
20
21 et al. (2007) consider standard questionnaire survey data as a snapshot in time. Hence,
22
23 the existing conditions are captured but these are not very suitable for representing
24
25 potential temporal variation (e.g., describing how the agents' decisions may change
26
27 owing to changes in their environment).
28
29

30
31 For these reasons, authors, such as Smajgl, et al. (2011), consider that a
32
33 standard questionnaire survey is useful for obtaining detailed descriptions of agent
34
35 attributes, but they suggest the use of other methods (e.g. interviews, field experiments,
36
37 role-playing games, or expert knowledge) to obtain data for understanding agent
38
39 behaviour. Our study will explore if the opportunity presented by the scenario-based
40
41 questionnaires, as described in the next section, can lead to reducing the drawbacks of
42
43 questionnaire survey data collection while, at the same time, retaining its advantages.
44
45
46
47

48 ***2.3. The use of scenario-based questionnaires in eliciting human decision rules***

49
50 In scientific research, scenario-based questionnaires are also known as vignette surveys.
51
52 We use the term “scenario-based questionnaire” because of its common use in the
53
54 supply chain research field encompassing our case study. In this survey method, a
55
56 scenario is a carefully constructed illustration of a person, object, or situation which
57
58 represents a systematic combination of characteristics. The wording of the scenario
59
60

1
2
3 embodies the factors to be tested and is experimentally controlled by the researcher
4
5 (Atzmüller and Steiner, 2010; "Encyclopedia of Survey Research Methods," 2008;
6
7 Lohfeld et al., 2012).
8
9

10 Some authors such as Atzmüller and Steiner (2010) consider the scenario-based
11
12 questionnaire as a powerful tool to elicit respondents' judgement. Consequently,
13
14 researchers have applied the scenario-based questionnaire in a variety of research fields.
15
16 For example, in operations management, it has been used to explore the factors
17
18 influencing the decision to outsource the manufacture of a component (Mantel,
19
20 Tatikonda, & Liao, 2006). Urda and Loch (2013) use scenarios to explain how emotions
21
22 and social preferences influence decision-making. Choo, Nag, & Xia (2015) investigate
23
24 how the style of executive problem solving influences knowledge accumulation and
25
26 manufacturing improvement. In operational research and decision science, Azadegan,
27
28 Golara, Kach, & Mousavi (2018) use scenarios to identify the drivers for managers to
29
30 increase their environmental investments. Similarly, Su, Chen, & Ro (2017) use a
31
32 scenario-based questionnaire to investigate the effects of individual negotiation styles
33
34 on the opportunism and compliance behaviours of buyers and suppliers. In computer
35
36 science, Jafarkarimi, Saadatdoost, Sim, & Hee (2016), employ scenario-based
37
38 questionnaires to study ethical dilemmas in using social networking sites. In the field of
39
40 R&D management, Cowlrick, Hedner, Wolf, Olausson, & Klofsten (2011) utilise it to
41
42 analyse entrepreneurial risk and attitude while there are plentiful examples from the
43
44 fields of health and psychology. All these studies affirm the benefits of a scenario-based
45
46 questionnaire in identifying the real actors' decision rules. However, none of these
47
48 studies uses the behaviours elicited to develop an ABMS model and this highlights a
49
50 clear research opportunity within the field of ABMS.
51
52
53
54
55
56
57
58
59
60

1
2
3 Typically, researchers derive the scenario-based questionnaire from theoretical
4 concepts (e.g., the theory of planned behaviour (Jafarkarimi, et al., 2016)). Using these
5 theories, they hypothesize the relationship between a series of factors with the
6 behaviour to be explored. Then they design the survey experiment by varying the level
7 of each factor (Atzmüller and Steiner, 2010) with full factorial design being the most
8 popular design of experiment. The researcher then converts each experimental set into
9 an illustration of a problem in which the respondents must make a decision (i.e., a
10 scenario). This step is important in that it can reduce the previously mentioned biases
11 arising from memory loss. These scenarios require the respondents to solve a current
12 and representative decision problem rather than try to recall a previous event (i.e., a
13 retrospective self-report). An earlier study has highlighted that scenarios designed using
14 real world situations allow the researcher to make generalizations or draw conclusions
15 about an individual's or a group's behaviours in reality (Cowlrick, et al., 2011). The
16 scenario format may also be more easily understood by respondents who do not have a
17 technical background when compared with asking them to confirm a decision tree or
18 flow chart.

19
20 We consider that there are similarities between the process for designing a
21 realistic ABMS and a scenario-based questionnaire. As explained earlier, an ABMS can
22 be developed theoretically or empirically. In principle, the initial (baseline) ABMS is a
23 collection of mechanisms and decision rules that are hypothesized by researchers
24 (Axelrod, 1997). The researchers then compare the ABMS's outputs with the real
25 phenomena. If there are similarities between the two, the researcher can claim that the
26 hypothesized mechanisms and decision rules are sufficient (though not necessarily
27 correct) to generate the observed real phenomena (Epstein, 1999). If we combine the
28 steps to design a scenario-based questionnaire with the ABMS then we can make a
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 stronger claim as to the model's validity because we can claim that the mechanisms and
4
5 decision rules that have been confirmed by respondents (rather than just hypothesized)
6
7 are sufficient to generate observable macro phenomena.
8
9

10
11 Consequently, in our study, we developed the scenarios for our survey using the
12
13 hypothesized decision rules created for the model (described further in Section 3) from
14
15 the findings of a literature review and then created a narrative for each scenario that was
16
17 adapted from the real world farmers' experience.
18
19

20 21 **3. Dairy Supply Chain in West Java and the Baseline Model**

22
23 ABMS has long been used to support policy making in the agri-food supply chain
24
25 (Utomo, et al., 2018). Policy makers in Indonesia have also realized the benefits of
26
27 simulation modelling in supporting policy making, including in the dairy supply chain
28
29 (Sunitiyoso, Wicaksono, Utomo, Putro, & Mangkusubroto, 2012). Currently, system
30
31 dynamics is the most widely used simulation methodology by the policy makers in
32
33 Indonesia (IAARD, 2012). For the dairy supply chain in particular, they have developed
34
35 sophisticated system dynamics models that can predict important parameters such as
36
37 milk production by considering a variety of factors such as changes in demographics
38
39 and macroeconomic conditions. Our full-scale ABMS aims to complement these models
40
41 by incorporating the real actors' individual behaviour. Such models will allow the
42
43 policy makers to analyse the effect of real actors' behaviours on the whole system
44
45 performance, and to design interventions that can encourage the real actors to adopt
46
47 more preferable behaviours. The model presented hereafter is a subset of our full-scale
48
49 ABMS, and focuses on farmers' decision-making rules in selling and buying cows. This
50
51 is because it is not possible to describe both the full-scale model and the process to
52
53 design the scenario-based questionnaire in detail in one paper.
54
55
56
57
58
59
60

1
2
3 The typical dairy supply chain in Indonesia is composed of many tiers
4
5 comprising farmers, cooperatives (collector and handler), milk processing industries
6
7 (processors), retailers and consumers. In common with earlier studies, the number of
8
9 farmers is large while the number of processors is very small (Glock, 2012). Most
10
11 farmers are smallholders who own relatively little land, which is only sufficient to build
12
13 a pen for their cattle and achieve relatively low production levels. For reasons of
14
15 security, the pens are usually located next to the farmers' houses in the middle of
16
17 residential areas. The forage grows along the road and riverbanks. The farmers gather
18
19 the forage from outside of their village using carts or motorcycles because it is difficult
20
21 for them to herd their cattle through the residential area. In this sense, forage is a
22
23 common resource for all these farmers. Hence, in the situation where the forage
24
25 availability is low, the competition between farmers to obtain forage should become
26
27 more intense and we expect certain behaviours would emerge as described for other
28
29 common resources such as fisheries (Bravo, 2011; Morano, de Moraes, & Jacomossi,
30
31 2018). However, in our case, such an extreme scarcity has not been observed in recent
32
33 years and we did not include this scenario in our questionnaire. Indeed, one of the major
34
35 concerns for policy makers is declining farmer and cow populations.
36
37
38
39
40

41
42 In this supply chain, the milk produced by the farmers is collected and
43
44 transported to the milk processors by farmers' cooperatives. The role of a farmers'
45
46 cooperative is important because it is cheaper for the milk processing industries to buy
47
48 milk in large quantities and, also, because it is highly perishable, the milk must be
49
50 transported efficiently and refrigerated at all times (Glover, Champion, Daniels, &
51
52 Dainty, 2014; Manish and Sanjay, 2013), which is prohibitively expensive for the
53
54 smallholder farmers. However, the cooperative's decisions are not fully controlled by
55
56 the farmers. The cooperative also has external investors, shareholders and employs
57
58
59
60

1
2
3 professional managers and workers. Hence, the cooperative operates like an
4
5 independent company with smallholder farmers acting as suppliers who have little
6
7 influence on the cooperative's decisions.
8
9

10 Consequently, we modelled a dyadic interaction between smallholder farmers
11
12 and the cooperative in West Java using [ABMS](#). The dairy supply chain in the case
13
14 study area is one of the biggest in Indonesia and we considered it representative of other
15
16 dairy supply chains in the country. Furthermore, we believe the case of the dairy supply
17
18 chain to be suitable to demonstrate the benefits of a scenario-based questionnaire
19
20 because the smallholder farmers (i.e., the respondents in our study) usually control their
21
22 own decisions. Hence, the respondent's answer will correspond directly with the agent's
23
24 decision rules in the simulation. This is in contrast to supply chains featuring large
25
26 organisations in which the decisions are more likely to be made by a management team
27
28 or via group agreement.
29
30
31

32
33 To develop the base model, we followed the suggestions of (Gilbert, 2004) and
34
35 collated the relevant body of knowledge from previous studies. During this literature
36
37 review, we found two sets of models relevant to the dairy supply chain. The first set of
38
39 models (e.g., [Happe, Schnicke, Sahrbacher, & Kellermann \(2009\)](#), [Happe, et al. \(2011\)](#),
40
41 [Marohn et al. \(2013\)](#), [Quang, Schreinemachers, & Berger \(2014\)](#)) assumes that farmers
42
43 have a land endowment. They maximize their income by allocating their land to
44
45 produce multiple crops. If the farmers decide to produce milk, then they allocate some
46
47 of their land to grow the forage. The second set of models comprise grazing models
48
49 (e.g., [Gross, McAllister, Abel, Smith, & Maru \(2006\)](#), [Boone et al. \(2011\)](#), [Martin,](#)
50
51 [Linstädter, Frank, & Müller \(2016\)](#), [Rasch, Heckelei, Oomen, & Naumann \(2016\)](#),
52
53 [Rasch, Heckelei, Storm, Oomen, & Naumann \(2017\)](#)) in which the farmers herd their
54
55 livestock to a common source of forage (i.e., the rangeland). In our case study area, the
56
57
58
59
60

1
2
3 farmers also mainly rely on their surrounding environment as a common source of
4
5 forage. Thus, we considered the second set of models to be more suitable as the
6
7 foundation for our base model. However, the farmers in our case need to transport
8
9 forage for their cattle while the cattle do not move at all. This introduced more
10
11 constraints into our modelling such as labour, working hours and transport capacity.
12
13

14
15 The focus of this paper is not to describe the base model development in detail
16
17 but to demonstrate how a scenario-based questionnaire is used to elicit key farmers'
18
19 decision rules to be implemented in the model. Consequently, we have summarised the
20
21 mechanism used in the base model, together with the main literature we have collated to
22
23 develop the base ABMS and our scenario-based questionnaire in Figure 1. A more
24
25 detailed description of each mechanism and the model is provided in the supplementary
26
27 materials.
28
29

30
31
32 Figure 1. Flowchart of the base ABMS accompanied with the main literature used to
33
34 develop each module

35 36 **4. Survey Instrument Design and Survey Process**

37 38 *4.1. Process to design the survey instrument*

39
40
41 The purpose of our survey was the collection of data that we could use to parameterise
42
43 our ABMS and elicit decision rules of the farmer agents in the model. The cooperative
44
45 was excluded from the survey because its decisions are made by many decision makers
46
47 collectively. Figure 2 describes the process we adopted to design our survey instrument.
48
49 This figure shows that questions aim to parameterise our ABMS are grouped in Part 1
50
51 of the survey instrument, while scenarios to elicit the agents' decision rules are grouped
52
53 in Part 2.
54
55
56
57
58
59
60

1
2
3 Figure 2. Flow chart of the process to develop the survey instrument
4

5 To develop the questions in Part 1, we began by listing the parameters used in
6
7 the base ABMS. These parameters included demographic (e.g., age and education),
8
9 socio-economic (e.g., income and off-farm jobs) and technical factors (e.g., cattle
10
11 ownership and cow productivity). The complete list of questions is given in the
12
13 supplementary materials.
14
15

16 The main contribution of this paper is in the development of scenarios in Part 2
17
18 of our survey instrument. In this particular case, we want to elicit farmers' buying and
19
20 selling decisions, which directly affect the size of the cattle and cow population. In
21
22 developing our scenario-based questionnaire, we were assisted by a number of domain
23
24 experts. The domain experts comprised lecturers and graduates from the Animal
25
26 Husbandry Department of Padjadjaran University and an experienced farmer. They
27
28 were chosen because they had experience in interacting with the farmers in the case
29
30 study area.
31
32
33

34 Once we were clear about the decision rules to be elicited, the next phase was to
35
36 list all the possible actions that can be taken by the farmers for each decision. The
37
38 purpose of the scenarios was to elicit the decision rules so it was important that this list
39
40 included not only actions that can be taken by the farmer agent in the base model, but
41
42 also other actions that may be performed in the real world. We also provided an option
43
44 where the respondents could explain actions that were not represented by other options.
45
46
47

48 In the third phase, we listed all parameters and information considered by the
49
50 agent to select its action in the base ABMS (decision parameters). We then determined
51
52 the range of decision parameter values that would be used to make variations of a
53
54 scenario. It is important that the decision parameter range included all values that can
55
56 occur in the simulation and the real world (*i.e.*, collectively exhaustive). This was done
57
58 to avoid bias owing to extrapolation (*i.e.*, when a decision parameter value that occurs
59
60

1
2
3 in the simulation goes beyond the range of data obtained from the respondent). If this
4 happens then the agents' decision rules in the simulation are no longer representative of
5 the real world actor. For example, in this study, we set the cattle mortality range in our
6 scenario between 0% - 100%. We also used information from the domain experts'
7 observations to establish these ranges, especially when the decision parameter
8 distribution is a priori unknown. For example, the domain experts observed that there
9 are farmers who start to sell their cattle when experiencing forage deficit for a week.
10 However, they also observed that some farmers will retain their cattle for two months
11 even though they are facing a forage deficit. Based on this information, we set the range
12 of farmer's forecast horizon in our scenario between one week and two months. If a
13 decision parameter proved to be significant while it was not possible to specify a
14 collectively exhaustive range for it, then a special error message would be created to
15 warn when its value in the simulation violates the data boundary. The corresponding run
16 should be then excluded from further analysis because it may contain bias.

17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The next phase was to combine these actions and decision parameters with a story to develop each scenario. This scenario guides the respondents to choose their actions by considering the given decision parameters. In this phase, we asked the domain experts to retell real farmers' experiences that they have observed. We then used the minimum, maximum and mid value of each decision parameter range to vary one scenario into several sub-scenarios using factorial design. Presenting several scenario variants is important to identify the sensitivity of a real actor's actions toward the changes in decision parameter value.

Finally, we translated the questions in Part 1 and Part 2 into local language and terminologies. We also used traditional measurement units in all of the survey

1
2
3 instrument questions to ensure that the respondents could understand all questions
4
5 easily.
6

7
8 We then asked the domain experts to validate our survey instrument through a
9
10 pilot test. There were several objectives of this pilot test. Firstly, it aimed to minimize
11
12 errors and ambiguity by asking the domain experts to propose revisions to the questions
13
14 or scenarios that were ambiguous or difficult to be understood. Secondly, the pilot
15
16 testing aimed to ensure that respondents' behaviours were sensitive to the scenarios
17
18 presented while keeping the questionnaire as short as possible. This was important
19
20 because the factorial design we used initially resulted in a massive number of scenarios
21
22 in the initial survey instrument design. We asked the domain experts to suggest new
23
24 parameter values if they felt that respondents' behaviour might not be sensitive towards
25
26 the decision parameter values presented. The domain experts could also propose new
27
28 action options and decision parameters for a scenario. The proposed action would be
29
30 considered in the revised survey instrument if it was mutually exclusive to the existing
31
32 options and it was plausible for a real farmer. If two adjacent sub-scenarios were
33
34 considered too similar and had no effect on the decision then the domain experts could
35
36 propose the elimination of one of the scenarios. When changing the decision parameter
37
38 values as well as eliminating a sub-scenario, it was necessary to keep the combination
39
40 of decision parameter value across all sub-scenarios collectively exhausted. By using
41
42 the domain experts' suggestions, we improved the survey instrument design over the
43
44 course of three iterations.
45
46
47
48
49
50

51 52 53 ***4.2. Survey and analysis process***

54
55 The full scale survey was carried out from 1st to 31st of August 2016. The respondents
56
57 comprised 153 farmer households located in 19 villages in the West Java area. The
58
59 respondents were identified from a database of the dairy cooperative members. Between
60

1
2
3 5 and 10 respondents were taken randomly from each village. The number of
4
5 respondents taken from a village depended on the proportion of the number of dairy
6
7 farmers in the particular village to all cooperative's members.
8
9

10 The scenario-based questionnaire is generally more complex than a standard
11 closed question questionnaire so each respondent was accompanied when completing it.
12
13 This is also important to ensure that respondents can interpret the survey instrument
14 correctly. The survey was conducted from house to house in the evening after the
15 respondents had finished all of their daily activities to ensure there were minimal
16 distractions for the respondents. We used hard copy format and on average, each
17 respondent required two hours to complete all the survey questions. Randomly, we
18 interviewed several respondents after they completed all the questions in the survey
19 instrument. In these interviews, we asked about their perceptions regarding the survey
20 instrument that we used. If they have participated in similar surveys (*e.g.*, agricultural
21 census), we also asked them to compare their experiences in responding to our scenario-
22 based questionnaire and a standard questionnaire.
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37

38 After the survey was completed, the respondent's responses were converted into
39 an electronic format using Microsoft Excel and SPSS. Codes were used to record
40 responses to closed answer questions in part 1 and part 2. If the respondent gave open
41 answers, the respondent's answers were then transcribed as sentences. This transcript
42 was then read in turn by the research team. The research team then agreed whether the
43 respondent's answer could be classified into one of the codes used in the closed answer
44 option.
45
46
47
48
49
50
51
52

53 Similarly for open answer responses, the respondent's response to interview
54 questions was transcribed into sentences. Each transcript was read in turn by the
55 research team. The research team then classified the respondents' response, for
56
57
58
59
60

1
2
3 example, as to whether the respondent can easily understand the survey instrument and
4
5 whether the respondent prefers the scenario-based questionnaire to the standard
6
7 questionnaire.
8
9

10 11 **5. Survey Findings**

12
13
14 Part 1 of the questionnaire shows that our respondents are quite homogeneous. That is,
15
16 98% of them are smallholders who have 7 or less cows, 85% of them have less than 600
17
18 m² of land, and 85% of them have been in dairy farming for less than 20 years. We also
19
20 obtained the parameter values and distributions that would be used in the simulation
21
22 model (Table 1).
23
24

25
26
27 Table 1. Simulation model parameters.
28
29

30
31 Part 2 of the questionnaire (i.e., the scenarios) is used to elicit farmers' buying
32
33 and selling decisions. These decisions are important because the objective of our model
34
35 is to predict the impact of these decisions on the volume of milk production and the size
36
37 of the cattle and cow population.
38
39

40 41 **5.1. Buying decision rule**

42
43
44 Gross, et al. (2006) suggest that the farmers' willingness to buy new cows is influenced
45
46 by the excess forage they obtain. Our domain experts who are mostly experts in animal
47
48 husbandry, suggested that higher milk price should also be considered as it has been
49
50 discussed in animal husbandry literature (e.g., Nicholson, Thornton, & Muinga (2004)).
51
52 Hence, we combined these two factors to represent the farmers' buying decision rule.
53
54
55 This combination produced four sub-scenarios that would be used to test whether those
56
57 factors were significant (Appendix, Scenario 1). In each sub-scenario, we asked the
58
59 respondent to state how many cows they are willing to buy assuming that they have
60

1
2
3 sufficient money.
4
5

6 We used regression analysis to estimate the number of cows that a farmer will
7 buy based on the two factors. Owing to the high skewness in the excess forage data, we
8 transformed it using a square root function to obtain a better fit. The regression model
9 (equation 1) shows that the only significant predictor is the square root of additional
10 forage obtained and it explained 14.7% of the variation in the number of cows a farmer
11 wanted to buy. The low coefficient value indicates that the farmers are risk averse. The
12 fact that milk price is not significant indicates that the farmers in our case study have
13 different behaviour from what has been discussed in the previous studies related to
14 animal husbandry. The complete statistical analysis is available from the supplementary
15 materials. The buying decision rule that we used to revise our baseline ABMS is shown
16 in Figure 3.
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31

$$32 \quad Add_{Cow} = -1.603 + 0.095Add_Forage^{\frac{1}{2}} \quad (1)$$

33
34
35
36
37 Figure 3. Buying decision rule
38

39 **5.2. Selling Decision Rule**

40
41 The literature suggests that the smallholder farmers' decision to sell their cows is
42 influenced by the forage deficit or financial problem they are experiencing (Boone, et
43 al., 2011; Gross, et al., 2006). Our domain experts suggested that when farmers sell
44 their cows due to forage deficit, it is mainly because of its impact on their cattle's
45 health. The farmers obtain their cattle's health information from a veterinarian who
46 visits them weekly. Based on this suggestion, we added the information provided by the
47 veterinarians concerning the cattle's health condition to the scenario. The scenario is
48 presented in the appendix (Scenario 2).
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 To test the influence of information provided by the veterinarians on the
4 probability of selling, we used logistic regression. The regression model is shown in
5 Equation 2. The model is significant and the Nagelkerke pseudo R^2 shows that the
6 model can predict 75% of respondents' response. The coefficient shows that as the
7 likelihood for the cow to become sick and die increases, the more the farmers choose to
8 sell the cow. We present the complete statistical analysis result in the supplementary
9 materials. The selling decision due to forage deficit is shown in the left flowchart on
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

$$\ln \frac{P_{sell}}{1 - P_{sell}} = 11.442P_{die} - 6.342 \quad (2)$$

Figure 4. Selling decision rules

When farmers are forced to sell cows owing to a financial problem, the literature typically assume that farmers prioritize the sale of the oldest cows first (e.g., Boone, et al. (2011)). During the pilot testing, our domain experts proposed other factors that might be considered by the farmers, namely: cow fertility and whether it is pregnant or not. Hence, we also incorporated these factors into our scenario-based questionnaire (Appendix, Scenario 3). In each scenario, we asked the farmers to compare two cows with different characteristics. We then asked them to choose which cow they preferred to sell. From this pairwise comparison, the research team then helped the respondents to order their preference from 1 (the most preferred) to 8 (the least preferred).

We used regression analysis to describe the farmers' preference. The regression model shows that all three factors are significant predictors of farmers' preference and the model R^2 is 97.8%. We present the complete statistical analysis result in the supplementary materials. Equation 3 shows that age becomes the first criteria in farmer

1
2
3 selection process (*Young* = 1 for young cows and 0 for old cows) followed by
4
5 pregnancy (*NotPregnant* = 1 for not pregnant cows and 0 for pregnant cows) and
6
7 fertility factors (*Fertile* = 1 for high fertility and 0 for low fertility), respectively. The
8
9 farmers place higher priority on selling a cow that is older, with low fertility, and not
10
11 pregnant. The selling decision due to financial problem is shown in the right flowchart
12
13 on Figure 4.
14
15

$$16 \text{ Priority} = 2.77 + 4 * \text{Young} + 1.22 * \text{Fertile} - 1.75 * \text{Not Pregnant} \quad (3)$$

20 6. Discussion

23 6.1. Respondents' perception toward the scenario-based questionnaire

24
25 Our experience during the data collection shows that most of the respondents could
26
27 understand the scenarios that were presented. During the data collection, there were a
28
29 few respondents who found it difficult to understand the scenarios. In these instances,
30
31 we asked them to express their understanding of the scenario presented. If the
32
33 respondent's understanding is not much different from the scenario's intention, then we
34
35 ask them to answer based on their understanding. Otherwise, we allow them to give no
36
37 response to the scenario (this only happens once during our data collection).
38
39
40
41

42
43 In addition, there were also some respondents who used the open answer option.
44
45 However, in our case study, we were always able to map their answer to one of the
46
47 predefined options. For example, some respondents mentioned that when they could not
48
49 get sufficient forage, they would look for alternatives to forage. In principle, this answer
50
51 shows that they prefer to retain the cows they own. Supposing in another case study that
52
53 the researcher cannot map the respondents' answer to the predefined options, we
54
55 propose to apply coding techniques to the respondents' answers. Indeed, the
56
57
58
59
60

1
2
3 applicability of the coding techniques depends on the number of respondents whose
4 answers cannot be accommodated by the predefined options.
5
6

7
8 Most of our respondents had taken part in previous studies that used
9
10 questionnaires as a data collection instrument. [One example of these studies is the](#)
11 [agricultural census conducted annually by the Indonesian Statistical Bureau](#). Therefore,
12 they had enough experience and knowledge to compare the benefits of the scenario-
13 based questionnaire that we used with the standard questionnaire. To reveal their
14 perception toward our design compared to the design in the previous studies, we
15 observed how our respondents react to the scenarios presented and conducted a short
16 interview with some of them after they completed the questionnaire.
17
18
19
20
21
22
23
24
25

26 More than 80% of the interviewees felt that they could understand the scenarios
27 presented because they were written in their daily language and terminology. According
28 to the interviewees, this questionnaire was different to the questionnaires in the previous
29 studies. In the previous surveys, it was difficult for them to imagine how the data would
30 be used and how the research outcome would be beneficial for them (partly because the
31 surveys often used technical terms and concepts that [are more familiar to the](#)
32 [academics](#)). In contrast, some of the interviewees could guess how the data from the
33 scenario-based questionnaire could be used to select interventions that might help them.
34 For example, one of the interviewees said that “If the government or cooperative know
35 that we decide to sell our cows because it is very difficult to collect sufficient forage,
36 then they could help us to import forage from other regions”. This finding shows that
37 we can identify key decision factors (i.e. the importance of forage) during the data
38 collection which help use devise potential policy interventions.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55

56 [The respondents’ perception during this interview provide the first insight](#)
57 [regarding the benefits of a scenario-based questionnaire compared to a standard](#)
58
59
60

1
2
3 questionnaire survey, namely: the concepts incorporated in a scenario-based
4
5 questionnaire can be more meaningful for the respondents. Yang and Gilbert (2008)
6
7 suggest that one of the differences between qualitative data and quantitative data relates
8
9 to how meaningful the concepts used are for the real world actors. Concepts used in
10
11 qualitative data collection are usually more meaningful for real world actors than the
12
13 concepts used in quantitative data collection, such as in a questionnaire survey. D. T.
14
15 Robinson, et al. (2007) consider this as one of the disadvantages of the survey as a data
16
17 collection methodology in ABMS because the respondents' might give responses to
18
19 concepts that are meaningless for them and hence bias the survey results. In common
20
21 with a standard questionnaire survey, our base ABMS and scenario-based questionnaire
22
23 were designed based on the previous literature. However, interviews with the
24
25 respondents indicate that the use of scenarios can help them to make the concepts used
26
27 in the modelling more meaningful. In addition, providing open answer options gives the
28
29 respondents opportunities to express their views on how they make the decisions in
30
31 reality. As a result, the data from a scenario-based questionnaire, which is designed with
32
33 sufficient pilot testing, can be more meaningful for the respondents.
34
35
36
37
38
39

40 The interviewees also found that the scenarios had occurred or were very likely
41
42 to occur in the real world. Those who ever faced similar situations claimed that their
43
44 responses to our questionnaire were similar to their actual actions back then. Those who
45
46 had never faced similar situations claimed that it was very likely that they would take
47
48 similar actions to their responses in the questionnaire. They also considered this design
49
50 to more beneficial for them because it stimulated them to think about their action if they
51
52 were to face a similar real scenario in the future.
53
54
55

56 The interviewees' responses provide the second insight regarding the usefulness
57
58 of a scenario-based questionnaire, namely: a scenario-based questionnaire can identify
59
60

1
2
3 how actors react to new scenarios. As mentioned earlier, the data obtained by a standard
4 questionnaire survey are mostly snapshots in time. Consequently, D. T. Robinson, et al.
5 (2007) suggest that the survey method is good for capturing the existing condition but
6 not very suitable for representing temporal variation. Longitudinal surveys are effective
7 in capturing temporal variation but this option can be expensive and is not always
8 feasible within the constraints of a research project. The interviewees reported that the
9 scenarios used in our study could help them to think about the actions they would take
10 in situations they had not yet experienced. This suggests that, though the scenario-based
11 questionnaire survey remains as a snapshot in time, we can still obtain indications of
12 how the real actors will choose their actions in possible future situations. Additionally,
13 these interviews provided an additional form of face-validation that gave us more
14 confidence that the decision rules revealed by the respondents reflected what they
15 actually do.

34 ***6.2. The benefits of a scenario-based questionnaire in improving ABMS validity***

35 The discussion in the previous section shows the benefits of scenario-based
36 questionnaire in eliciting decision rules that can be used in an ABMS model. Another
37 benefit of scenario-based questionnaire is that it allows us to elicit decision rules
38 empirically from the farmers which gives us more confidence in the representativeness of
39 the behaviours of agents in our model. This is known as micro-validity (Takadama,
40 Kawai, & Koyama, 2008).

41 However, model validation in ABMS needs to be done at least in two levels:
42 **micro-validation and macro-validation**. Micro-validation evaluates whether the
43 behaviours of individual agents in an ABMS model correspond with the observed
44 behaviours of real-world actors. Macro-validation evaluates whether the behaviours
45 emerging from the interactions between agents in an ABMS model corresponds with the
46

observed system level behaviour in the real world. Both micro- and macro-validation are especially important if researchers seek truer representations of human behaviour in studies that uses [ABMS \(Macal, 2016\)](#).

In this section, we investigate whether elicited decision rules can improve the model's macro-validity. [The validity concept that we use is operational validity, which considers the match between the ABMS output to the real data \(Sargent, 2013\)](#). The system-level outputs that we use in macro-validation are cattle population, cow population and milk production as these are considered to be important by both the government and cooperative when recording their statistics (Table 2).

Table 2. Cattle population, cow population and average daily milk production in Pangalengan West Java 2010-2012 (KPBS, 2016).

In Table 3, we compare the macro-validity of the base model (M0) with the empirical models (M1 – M3). The empirical models use one or more elicited decision rules (i.e. buying and selling) obtained from the scenario-based questionnaire. We use the mean error estimation to measure the magnitude of model output deviations from the real data. To estimate the mean error, we measured the difference between model outputs at the end of each simulation year and the real data, from 2010 until 2012 (i.e., $Error_i = Data_i - Simulation_i$ where $i = 2010 \dots 2012$). We then computed the mean error (ME) from 2011 to 2012 (i.e. $ME = \sum_{2011}^{2012} Error_i / 2$). Table 3 shows the average (\overline{ME}) and standard deviation (S_{ME}). A t-test was then carried out to infer whether, in the long run, the model's average ME is zero. The two-tailed significance (sig. column) of the t-test at 95% confidence level is also presented. A lower $|\overline{ME}|$ value indicates that on average the model output is closer to the real data. While, a significance value higher than 5% indicates that we fail to reject the null hypothesis that the simulation output

1
2
3 reflects the real-world data (i.e., a valid model). The result shows that the base model is
4
5 only valid for one output, i.e. cow population. M1-M3 are valid for at least two outputs,
6
7 indicating models with empirical decision rules are better in terms of improved model
8
9 macro-validity.
10
11
12

13 Table 3. The macro-validation of the ABS model.

14
15 This finding shows that scenario-based questionnaire can be used to increase the
16
17 micro and macro validity of our ABS model. Macal (2016) identify model
18
19 parameterisation (including parameterising the decision rules) and validation based on
20
21 empirical data as one of the key challenges in ABMS research. Hence, this paper shows
22
23 that scenario-based questionnaire provides us with a tool to collect empirical data for
24
25 ABMS model parameterisation and validation.
26
27
28
29

30 *6.3. Other benefits and weaknesses of a scenario-based questionnaire*

31
32 This section discusses the comparison of our experiences in developing and deploying a
33
34 scenario-based questionnaire with the views expressed in the previous literature. We
35
36 consider that there are other benefits of a scenario-based questionnaire, other than those
37
38 discussed in the two previous sections, namely:
39
40

- 41 1. *A scenario-based questionnaire enables the clarification of the context of the*
42 *agents' decisions.* According to Yang and Gilbert (2008), surveys that are
43
44 usually used to collect quantitative data place less emphasis on context (i.e.,
45
46 whether and how a decision rule is activated by considering an agent's current
47
48 state and environment). Furthermore, An (2012) observes that the statistical
49
50 methods commonly used to analyse survey data are often problematic in
51
52 providing insight into an agent's motive, incentive and preferences when making
53
54 a decision. Our experience shows that, with careful pilot testing, it is possible to
55
56 identify the context of a decision rule using a scenario-based questionnaire. For
57
58
59
60

1
2
3 example, we identified that cattle health conditions rather than forage shortage
4 trigger the farmers' selling decision rule. These health condition scenarios were
5 proposed by our pilot testing respondents. By applying statistical analysis to the
6 survey data, we were also able to identify farmers' preferences when selecting
7 the cow to be sold.
8
9

- 10
11
12
13
14
15 2. *A variety of established statistical techniques can be used to analyse the data*
16 *obtained from a scenario-based questionnaire to create decision rules.* For
17 example, in our study, the selling decision rule is binary and we used
18 multinomial logistic regression to extract the decision rule and, alternatively, we
19 could use techniques such as curve fitting. It is also possible to incorporate the
20 effect of agent heterogeneity in the decision rule as suggested by D. T.
21 Robinson, et al. (2007). For example, this can be achieved by clustering agents'
22 attributes (e.g., based on demography and socioeconomic parameters, as by
23 Valbuena, et al. (2008)) or by using these attributes as dummy and control
24 variables in a regression model.
25
26
27
28
29
30
31
32
33
34
35
36

37 Nevertheless, the scenario-based questionnaire in this study also inherits the weaknesses
38 of the survey method. For example, our survey assumes that the head of the farmer
39 household is the sole decision maker in the family (D. T. Robinson, et al., 2007). In
40 reality, each family member may contribute opinions and thoughts when the head of the
41 household make a decision. Also, we rely on statistical techniques to analyse the data
42 and these techniques rely upon many structural and technical assumptions (D. T.
43 Robinson, et al., 2007). Similarly, extrapolation based upon statistical analyses of
44 survey data needs care. Relationships derived from the analyses of survey data can be
45 good at estimating values within the data range (i.e., interpolate). However, when the
46 simulation is running there is potential for the variable values to exceed the range of
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 empirical data. In this case, the decision rules derived from the survey are used for
4
5 extrapolation. When this happens, the decision rule in the ABMS is not representative
6
7 of the actual agents even if, on aggregate, our simulation result is valid when compared
8
9 to the real data. We attempted to minimise this potential bias by defining collectively
10
11 exhaustive parameter ranges to be used during the scenario design process. There were
12
13 several parameters whose range was a priori unknown (e.g., how long the farmers
14
15 experience forage shortage before they eventually decide to sell their cows) but,
16
17 fortunately, these parameters did not significantly affect the farmers' decisions. If these
18
19 parameters were significant then we could have avoided the potential bias by excluding
20
21 simulation runs in which these parameters' values exceeded the data range.
22
23
24
25
26
27

28 **7. Conclusion**

29
30 The main contribution of this paper is the detailed discussion of steps to design a
31
32 scenario-based questionnaire. We have tested the usefulness of scenario-based
33
34 questionnaires in the case of a dairy supply chain and recorded the respondents'
35
36 perceptions of the survey. We have shown that the data obtained through the survey is
37
38 useful for parameterising simulation inputs and improving the micro- and macro-
39
40 validity of the ABMS model. Scenario-based questionnaire is particularly useful to
41
42 elicit behaviours where the respondent's answer corresponds directly with the agent's
43
44 decision rules in the simulation, for example the smallholder farmers in our case study
45
46 usually control their own decisions (in contrast to supply chains featuring large
47
48 organisations in which the decisions are more likely to be made by a management team
49
50 or via group agreement). Hence, we believe our method can be applied to other study on
51
52 supply chains with similar characteristics in which ABMS models is used.
53
54
55
56
57
58
59
60

References

- An, L. (2012). Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling*, 229, pp. 25-36.
- Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology*
- Axelrod, R. (1997). Advancing the art of simulation in the social sciences *Simulating social phenomena* (pp. 21-40): Springer.
- Azadegan, A., Golar, S., Kach, A., & Mousavi, N. (2018). Corporate environmental investments: A cross-national study on managerial decision making. *International Journal of Production Economics*, 199, pp. 47-64.
- Bahar, S. (2014). Produktivitas hijauan pakan untuk produksi sapi potong di Sulawesi Selatan. *JITV*, 19(3)
- Bankes, S. C. (2002). Agent-based modeling: A revolution? *Proceedings of the National Academy of Sciences*, 99(suppl 3), pp. 7199-7200.
- Barlas, P., Heavey, C., & Dagkakis, G. (2015). An open source tool for automated input data in simulation. *International Journal of Simulation Modelling (IJSIMM)*, 14(4)
- Boone, R. B., Galvin, K. A., BurnSilver, S. B., Thornton, P. K., Ojima, D. S., & Jawson, J. R. (2011). Using coupled simulation models to link pastoral decision making and ecosystem services. *Ecology and society*, 16(2)
- Bravo, G. (2011). Agents' beliefs and the evolution of institutions for common-pool resource management. *Rationality and Society*, 23(1), pp. 117-152.
- Bruch, E., & Atwell, J. (2015). Agent-based models in empirical social research. *Sociological methods & research*, 44(2), pp. 186-221.
- Choo, A. S., Nag, R., & Xia, Y. (2015). The role of executive problem solving in knowledge accumulation and manufacturing improvements. *Journal of operations management*, 36, pp. 63-74.
- Cowlrick, I., Hedner, T., Wolf, R., Olausson, M., & Klofsten, M. (2011). Decision-making in the pharmaceutical industry: analysis of entrepreneurial risk and attitude using uncertain information. *R&D Management*, 41(4), pp. 321-336.
- Eldabi, T., Irani, Z., Paul, R. J., & Love, P. E. (2002). Quantitative and qualitative decision-making methods in simulation modelling. *Management Decision*, 40(1), pp. 64-73.
- . Encyclopedia of Survey Research Methods. (2008). In P. J. Lavrakas (Ed.), *Encyclopedia of Survey Research Methods*. Thousand Oaks, California: SAGE Publishing.
- Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, 4(5), pp. 41-60.
- Gilbert, N. (2004). Agent-based social simulation: dealing with complexity. *The Complex Systems Network of Excellence*, 9(25), pp. 1-14.
- Glock, C. H. (2012). Coordination of a production network with a single buyer and multiple vendors. *International Journal of Production Economics*, 135(2), pp. 771-780.
- Glover, J. L., Champion, D., Daniels, K. J., & Dainty, A. J. (2014). An Institutional Theory perspective on sustainable practices across the dairy supply chain. *International Journal of Production Economics*, 152, pp. 102-111.
- Gross, J., McAllister, R. R., Abel, N., Smith, D. S., & Maru, Y. (2006). Australian rangelands as complex adaptive systems: a conceptual model and preliminary results. *Environmental Modelling & Software*, 21(9), pp. 1264-1272.

- 1
2
3 Hahn, H. A. (2013). The conundrum of verification and validation of social science-
4 based models. *Procedia Computer Science*, 16, pp. 878-887.
- 5 Happe, K., Hutchings, N., Dalgaard, T., & Kellerman, K. (2011). Modelling the
6 interactions between regional farming structure, nitrogen losses and
7 environmental regulation. *Agricultural Systems*, 104(3), pp. 281-291.
- 8 Happe, K., Schnicke, H., Sahrbacher, C., & Kellermann, K. (2009). Will they stay or
9 will they go? Simulating the dynamics of single-holder farms in a dualistic farm
10 structure in Slovakia. *Canadian Journal of Agricultural Economics/Revue*
11 *canadienne d'agroeconomie*, 57(4), pp. 497-511.
- 12 Heath, B., Hill, R., & Ciarallo, F. (2009). A survey of agent-based modeling practices
13 (January 1998 to July 2008). *Journal of Artificial Societies and Social*
14 *Simulation*, 12(4), p 9.
- 15 IAARD. (2012). *Kebijakan pencapaian swasembada dan swasembada berkelanjutan*
16 *lima komoditas utama pertanian melalui pendekatan sistem dinamik* Jakarta
17 Selatan: Indonesian Agency for Agricultural Research and Development.
- 18 Jafarkarimi, H., Saadatdoost, R., Sim, A. T. H., & Hee, J. M. (2016). Behavioral
19 intention in social networking sites ethical dilemmas: An extended model based
20 on theory of planned behavior. *Computers in Human Behavior*, 62, pp. 545-561.
- 21 Jager, W., & Janssen, M. (2002). *The need for and development of behaviourally*
22 *realistic agents*. International Workshop on Multi-Agent Systems and Agent-
23 Based Simulation.
- 24 Janssen, M., & Ostrom, E. (2006). Empirically based, agent-based models. *Ecology and*
25 *society*, 11(2)
- 26 KPBS. (2016). *Data Populasi dan Penghasilan Anggota KPBS* [Data set]. Retrieved
27
- 28 Lie, H., & Rich, K. (2016). Modeling dynamic processes in smallholder dairy value
29 chains in Nicaragua: a system dynamics approach.
- 30 Lie, H., Rich, K. M., & Burkart, S. (2017). Participatory system dynamics modelling for
31 dairy value chain development in Nicaragua. *Development in Practice*, 27(6),
32 pp. 785-800.
- 33 Lohfeld, L., Goldie, J., Schwartz, L., Eva, K., Cotton, P., Morrison, J., . . . Wood, T.
34 (2012). Testing the validity of a scenario-based questionnaire to assess the
35 ethical sensitivity of undergraduate medical students. *Medical teacher*, 34(8),
36 pp. 635-642.
- 37 Macal, C. M. (2016). Everything you need to know about agent-based modelling and
38 simulation. *Journal of Simulation*, 10(2), pp. 144-156.
- 39 Manish, S., & Sanjay, J. (2013). Agri-fresh produce supply chain management: a state-
40 of-the-art literature review. *International Journal of Operations & Production*
41 *Management*, 33(2), pp. 114-158.
- 42 Mantel, S. P., Tatikonda, M. V., & Liao, Y. (2006). A behavioral study of supply
43 manager decision-making: Factors influencing make versus buy evaluation.
44 *Journal of operations management*, 24(6), pp. 822-838.
- 45 Marohn, C., Schreinemachers, P., Quang, D. V., Berger, T., Siripalangkanont, P.,
46 Nguyen, T. T., & Cadisch, G. (2013). A software coupling approach to assess
47 low-cost soil conservation strategies for highland agriculture in Vietnam.
48 *Environmental Modelling & Software*, 45, pp. 116-128.
- 49 Martin, R., Linstädter, A., Frank, K., & Müller, B. (2016). Livelihood security in face of
50 drought—assessing the vulnerability of pastoral households. *Environmental*
51 *Modelling & Software*, 75, pp. 414-423.
- 52 Morano, R. S., de Moraes, E. A., & Jacomossi, R. R. (2018). Can small groups avoid
53 the tragedy of the commons? *AI & SOCIETY*, 33(1), pp. 71-80.
- 54
55
56
57
58
59
60

- 1
2
3 Morgan, F., Brown, P., & Daigneault, A. (2015). Simulation vs. definition: Differing
4 approaches to setting probabilities for agent behaviour. *Land*, 4(4), pp. 914-937.
- 5 Moss, S. (2008). Alternative approaches to the empirical validation of agent-based
6 models. *Journal of Artificial Societies and Social Simulation*, 11(1), p 5.
- 7
8 Nicholson, C. F., Thornton, P. K., & Muinga, R. W. (2004). Household-level Impacts of
9 Dairy Cow Ownership in Coastal Kenya. *Journal of Agricultural Economics*,
10 55(2), pp. 175-195.
- 11 Onggo, B. S., & Hill, J. (2014). Data identification and data collection methods in
12 simulation: a case study at ORH Ltd. *Journal of Simulation*, 8(3), pp. 195-205.
- 13 Onggo, B. S., Hill, J., & Brooks, R. J. (2013). *A pilot survey on data identification and*
14 *collection in simulation projects*. European Simulation and Modeling
15 Conference.
- 16 Perera, T., & Liyanage, K. (2000). Methodology for rapid identification and collection
17 of input data in the simulation of manufacturing systems. *Simulation Practice*
18 *and Theory*, 7(7), pp. 645-656.
- 19 Quang, D. V., Schreinemachers, P., & Berger, T. (2014). Ex-ante assessment of soil
20 conservation methods in the uplands of Vietnam: An agent-based modeling
21 approach. *Agricultural Systems*, 123, pp. 108-119.
- 22 Rasch, S., Heckeley, T., Oomen, R., & Naumann, C. (2016). Cooperation and collapse
23 in a communal livestock production SES model—a case from South Africa.
24 *Environmental Modelling & Software*, 75, pp. 402-413.
- 25 Rasch, S., Heckeley, T., Storm, H., Oomen, R., & Naumann, C. (2017). Multi-scale
26 resilience of a communal rangeland system in South Africa. *Ecological*
27 *Economics*, 131, pp. 129-138.
- 28 Robinson, D. T., Brown, D. G., Parker, D. C., Schreinemachers, P., Janssen, M. A.,
29 Huigen, M., . . . Irwin, E. (2007). Comparison of empirical methods for building
30 agent-based models in land use science. *Journal of Land Use Science*, 2(1), pp.
31 31-55.
- 32 Robinson, S. (2004). *Simulation: the practice of model development and use*: Wiley
33 Chichester.
- 34 Sargent, R. G. (2013). Verification and validation of simulation models. *Journal of*
35 *Simulation*, 7(1), pp. 12-24.
- 36 Schutte, S. (2010). Optimization and falsification in empirical agent-based models.
37 *Journal of Artificial Societies and Social Simulation*, 13(1)
- 38 Skoogh, A., & Johansson, B. (2008). *A methodology for input data management in*
39 *discrete event simulation projects*. 2008 Winter Simulation Conference.
- 40 Smajgl, A., Brown, D. G., Valbuena, D., & Huigen, M. G. (2011). Empirical
41 characterisation of agent behaviours in socio-ecological systems. *Environmental*
42 *Modelling & Software*, 26(7), pp. 837-844.
- 43 Su, H.-C., Chen, Y.-S., & Ro, Y. K. (2017). Perception differences between buyer and
44 supplier: the effect of agent negotiation styles. *International Journal of*
45 *Production Research*, 55(20), pp. 6067-6083.
- 46 Sunitiyoso, Y., Wicaksono, A., Utomo, D. S., Putro, U. S., & Mangkusubroto, K.
47 (2012). Developing strategic initiatives through Triple Helix interactions:
48 Systems modelling for policy development. *Procedia-Social and Behavioral*
49 *Sciences*, 52, pp. 140-149.
- 50 Takadama, K., Kawai, T., & Koyama, Y. (2008). Micro-and macro-level validation in
51 agent-based simulation: reproduction of human-like behaviors and thinking in a
52 sequential bargaining game. *Journal of Artificial Societies and Social*
53 *Simulation*, 11(2), p 9.
- 54
55
56
57
58
59
60

- 1
2
3 Trybula, W. (1994). *Building simulation models without data*. Proceedings of IEEE
4 International Conference on Systems, Man and Cybernetics.
5 Urda, J., & Loch, C. H. (2013). Social preferences and emotions as regulators of
6 behavior in processes. *Journal of operations management*, 31(1-2), pp. 6-23.
7 Utomo, D. S., Onggo, B. S., & Eldridge, S. (2018). Applications of agent-based
8 modelling and simulation in the agri-food supply chains. *European Journal of*
9 *Operational Research*, 269(3), pp. 794-805.
10 Valbuena, D., Verburg, P. H., & Bregt, A. K. (2008). A method to define a typology for
11 agent-based analysis in regional land-use research. *Agriculture, Ecosystems &*
12 *Environment*, 128(1-2), pp. 27-36.
13 Yang, L., & Gilbert, N. (2008). Getting away from numbers: Using qualitative
14 observation for agent-based modeling. *Advances in complex systems*, 11(02), pp.
15 175-185.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

APPENDIX: Scenario-Based Questionnaire

Scenario 1: These scenarios are used to validate and calibrate buying decision rules

Scenario 1a: In the current condition in which you can collect forage (*respondent's answer to Part 1*) and milk price of (*respondent's answer to Part 1*), how many more cows do you want to buy, suppose you have enough money to buy the cows and to increase your pen capacity?

Scenario 1b: Please imagine a condition in which the forage availability has increased drastically. With the same amount of labour and time, you can collect twice as much forage as the forage you can collect at this time. However, the milk price you receive stays the same. If you have enough money to buy new cows and to increase your pen capacity, then how many new cows do you want to buy?

Scenario 1c: Suppose the forage availability stays the same but the milk price is double. If you have enough money to buy new cows and to increase your pen capacity, then how many new cows do you want to buy?

Scenario 1d: Please imagine a condition in which the forage availability has increased drastically. With the same amount of labour and time, you can collect twice as much forage as the forage you can collect at this time. In addition, the milk price is also double. If you have enough money to buy new cows and to increase your pen capacity, then how many new cows do you want to buy?

Scenario 2: These scenarios are used to validate and calibrate selling decision rules

Please imagine that you only have one cow. Unfortunately, you are facing drought in the last 7 days and during this period you can only satisfy 75% of the forage needed by your cow. When the veterinarian come for his regular visit, he tells you that there is

1
2
3 25% chance of your cow will be sick and die tomorrow. Soon after the veterinarian
4 leaves, you receive a call from a butcher, offering to buy your cow for 15 million. This
5 price is acceptable considering your cow live weight. If you accept the butcher's offer
6 while the veterinarian's prediction does not happen then you lose your potential future
7 income. On the other hand, if you decline this offer and the veterinarian's prediction
8 happen then you will not get anything. In this condition which action will you take?
9

10
11
12
13
14
15
16
17 (a) To sell your cow; (b) to retain your cow; (c) Other, please explain
18

19 Notes: For each respondent, we ask several sub-scenarios by varying the drought
20 period, forage sufficiency and cow mortality. The drought period variation is 7 days, 1
21 month, and 2 months. The forage sufficiency variation is 0%, 50% and 75%. The
22 probability to die variation is 0%, 25%, 50% and 75%. If it is difficult for the
23 respondent to imagine probability using percentage, then the information is rephrase
24 using odds (e.g., in one occasion your cow will die and in 3 occasions your cow can
25 survive).
26
27
28
29
30
31
32
33
34
35
36

37 ***Scenario 3: These scenarios are used to validate and calibrate cow selection***
38 ***decision rule***
39

40 Please imagine that you have only two cows. You are currently experiencing financial
41 difficulties and are unable to get help, hence you need to sell one of your cows. The
42 money from selling one of these cows can meet your current needs. The first cow is
43 young, currently, it is not pregnant but it can get pregnant easily when given artificial
44 insemination. Your second cow is old, from your record it is hard to get pregnant when
45 it is artificially inseminated, but currently it is pregnant. Which cow do you prefer to
46 sell?
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1. Flowchart of the base ABMS accompanied with the main literature used to develop each module

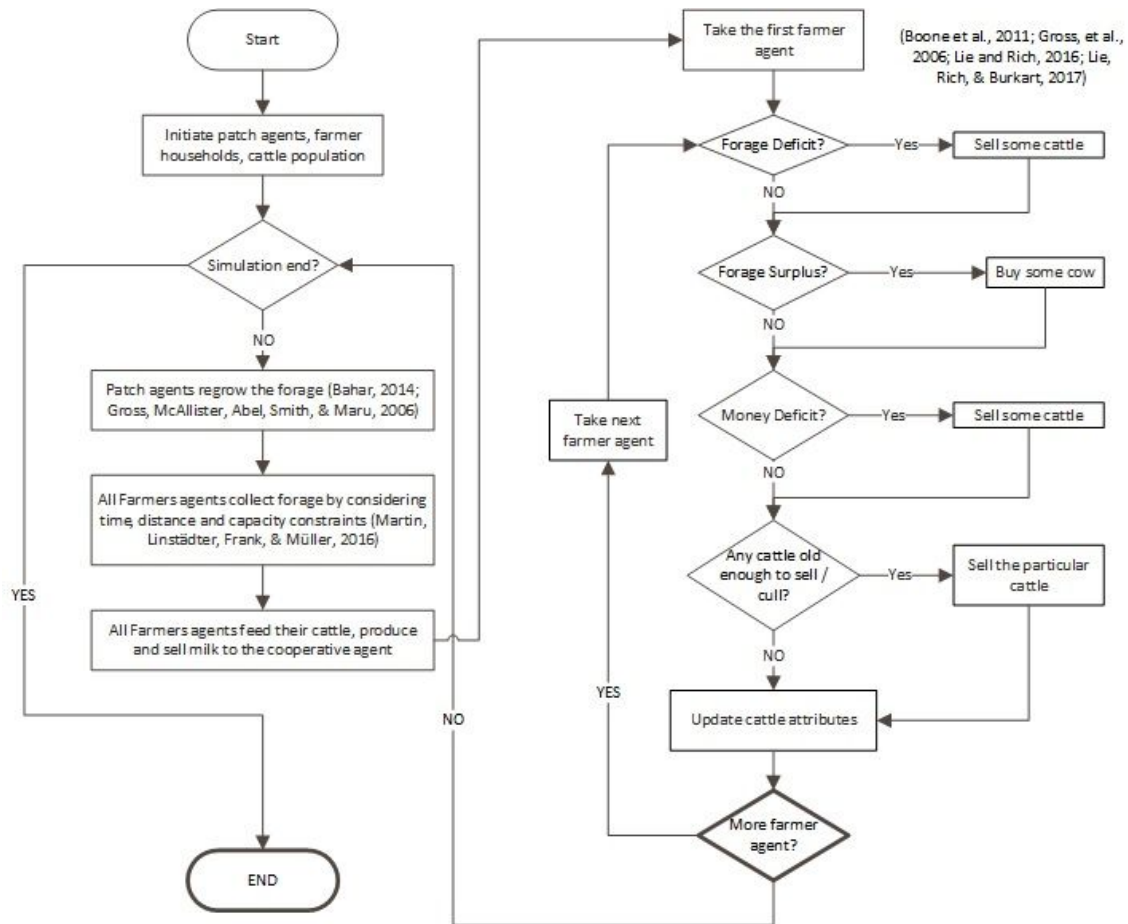
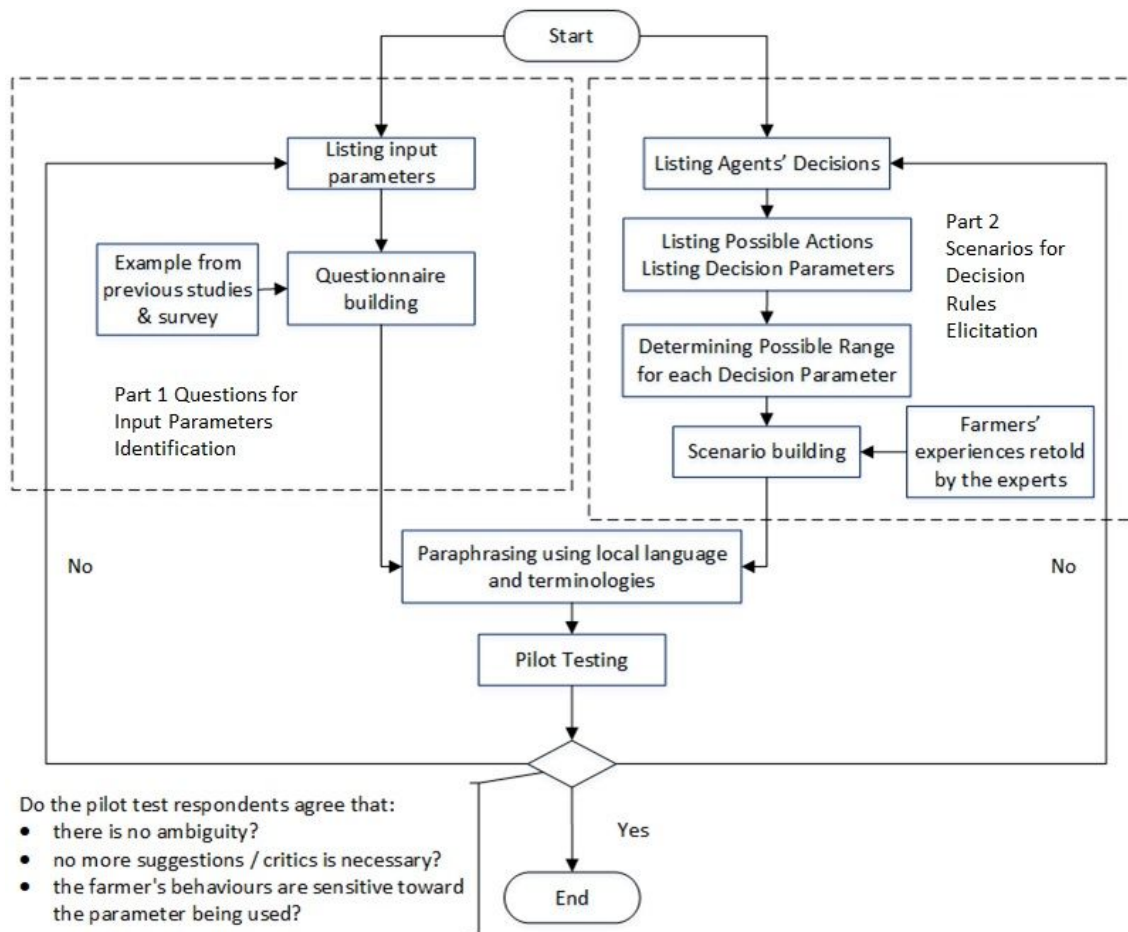


Figure 2. Flow chart of the process to develop survey instrument



www.Only

Figure 3. Buying decision rule

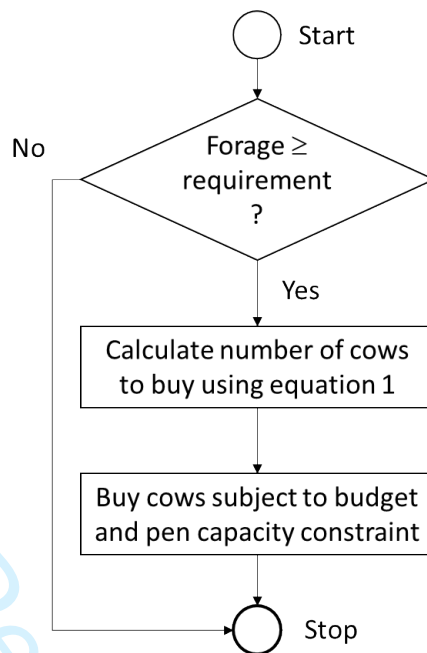


Figure 4. Selling decision rules

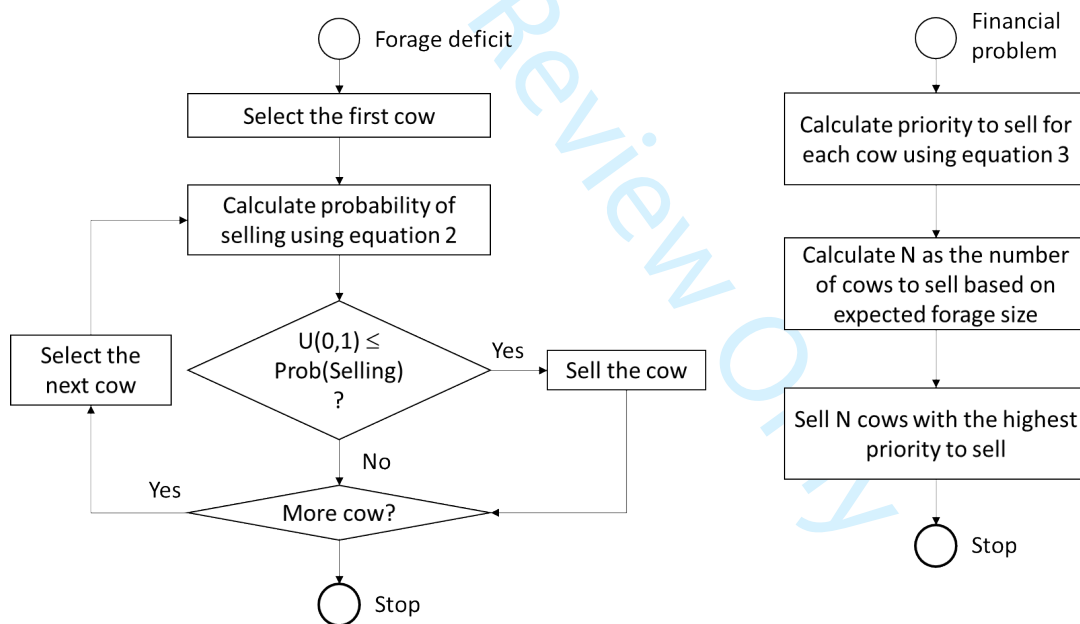


Table 1. Simulation model parameters.

Variable Name	Descriptive statistics					Distribution
	Min	Max	Mode	Mean	Std. Dev	
Agent attributes						
Farmer Age (years)	22	74	38	46.17	10.98	Triangular
Family Labour (person)	0	4	1	0.92		Binomial
Number of Cow (heads)	0	18	3	4.10		Poisson
Number of Bull (heads)	0	5	0	0.81		Poisson
Peak Milk Prod (litre)	10	35	20	20.81	19.35	Normal
Service per conception (times)	1	8	2	2.38		Poisson
Constants						
Cow Selling Price (millions Rp/head)				13.1		
Bull Selling Price (millions Rp/head)				16.4		
Heifer Buying Price (millions Rp/head)				9.6		
Minimum Milk Price (Rp/litre)				3,350		
Maximum Milk Price (Rp/litre)				5,200		
Additional Fodder Price (Rp/Kg)				2,400		

Table 2. Cattle population, cow population and average daily milk production in Pangalengan West Java 2010-2012 (KPBS, 2016).

Year	Cattle population (head)	Cow population (head)	Average daily Production (litre)
2010	21,322	21,083	159,333
2011	21,438	20,960	136,694
2012	22,366	22,073	138,904

Table 3. The macro-validation of the ABS model.

Model Name	Cattle Population			Cow Population			Daily Milk Production		
	\overline{ME}	S_{ME}	Sig.	\overline{ME}	S_{ME}	Sig.	\overline{ME}	S_{ME}	Sig.
M0 Base model	-2272.3	4395.4	0.02	-1443.1	4025.3	0.09	20600.2	13421.1	0.00
M1 Buying decision	-1494.8	5075.0	0.15	-876.3	4703.7	0.36	16811.3	15151.1	0.00
M2 Selling decisions	-1472.9	5104.2	0.16	-874.6	4586.1	0.35	6359.1	18349.6	0.10
M3 Buying & Selling decisions	-1504.0	5116.5	0.15	-904.4	4588.1	0.33	6118.6	18383.3	0.11

Eliciting Agents' Behaviour using Scenario-Based Questionnaire in Agent-Based Dairy Supply Chain Simulation [Supplementary Materials]

Appendix A: Base Model Description

The model aims to replicate the dynamics of milk production, cow population and cattle population in the case study area. Those variables are important for policy makers as indicated by the annual reporting of industry statistics. To produce these outputs, the model uses several inputs such as the initial number of farmer households, number of family labour, cattle ownership, and cow productivity. The values of these parameters are identified using the survey described in the paper.

There are three types of agent in the model, namely: a number of farmer households, a cooperative and forage patches. The farmer household's role is to produce and supply milk to the cooperative. The cooperative sets the milk price based on the milk quality and then sells the milk to the milk processing industry. The farmers interact with the patches whose main function is to provide forage for their cows. The conditions in the case study area are representative for the typical supply chain in Indonesia though the arrangement of the agents in the system may vary. The simulation operates on daily time step, although some processes occur on a monthly and annual schedule.

A.1 The patch agent

There are 306 patches in the model; each represents an area of one kilometre square. We define three types of patch: used patch, unused patch and forage patch. Used patches represent the land area that has been occupied by building, houses, roads, etc. Unused patches represent empty land areas that can be used to build new cattle pens but not to grow forage. Forage patches represent land areas that are overgrown with forage.

Every day the patches produce forage. Equation A.1 describes the amount of forage production (kg per km² per day).

$$\frac{dF}{dt} = \text{Min}((F_{max} - F_t - Fc_t), (F_t - Fc_t) * (1 + G)) \quad (\text{A.1})$$

F_{max} represents the maximum amount of forage (kg) per kilometre square area. There are various forage grass species in the case study area and the actual composition is unknown. However, Bahar (2014) estimates the forage productivity of various grass composition that can grow in a one kilometre square area in Indonesia is between 270 and 734 tonnes per km². Hence, in each run, the maximum amount of forage that can grow on a patch is randomized within this range using a uniform distribution. F_t is the initial forage level at day t and Fc_t is the amount of forage taken by the farmers on day t . G represents the forage growth rate with an average value of 1.1% per day (Bahar, 2014)

A.2 The farmer household agent

A farmer household agent consists of several family members who work together to rear cattle. Each farmer household has several attributes. Some of the farmer's attributes are modelled as variables (e.g., money, number of cattle, pen area and type of transportation mode). Other

farmer attributes are modelled as lists (e.g., family members' age, cattle gender, cattle age, the percentage of fodder fulfilment, services per conception and maximum milk production). Each element in the services per conception and maximum milk production list represents the fertility and the maximum milk that can be produced by each cow respectively. The elements in these lists only have a non-zero value for the cows.

In line with the previous studies, we assume that the farmers accumulate their assets (i.e., money and cattle) over time (Gross, McAllister, Abel, Smith, & Maru, 2006). Farmers' income comes from milk and cattle selling and they use their money to pay monthly living expenses. According to our experts, rearing dairy cattle is very time consuming so very few farmers have sources of income other than producing and selling milk. Therefore, we assume that farmers do not produce crops or have off-farm jobs. In common with earlier studies, we assume that farmers increase their assets by using strategies to collect forage, sell milk, sell cattle, buy cattle and expand pen area (Gross, et al., 2006).

A.2.1 The forage collection process

Every day farmers collect forage to feed their cattle. They scan forage patches around their house. The number of working hours and the speed of the transportation mode at their disposal limit the maximum distance they can travel. Each farmer household typically has 8 hours per day to collect forage during the period between the cooperative's milk collections (i.e., 7 am and 3 pm). In the case study area, the farmers collect forage on foot or by motorcycle or truck. In common with Martin, Linstädter, Frank, & Müller (2016), we assume that farmers prioritize the location with the highest forage level when choosing the location to collect forage. If there is more than one location with the highest forage level then farmers prioritize forage collection from the closest location to their house.

Having decided the location to collect forage, the agents move to the designated patch. Their travel time is subtracted from their remaining working hours. The amount of forage they can collect from the given patch is constrained by the patch's forage level, the amount of family labour, their remaining working hours and their transport capacity. Actual measurements regarding these variables were not available so we asked our expert to suggest reasonable approximations based on his experience. The expert suggested that each family labourer could harvest 40 kg of forage per hour. Furthermore, the expert suggested that they could carry 40 kg of forage per person per trip if they transported the forage on their back or using a cart, 60 kg of forage per trip if they used motorcycle and 600 kg of forage per trip if they used a truck.

A.2.2 The cattle feeding process

Farmer agents use the forage to feed their cattle. The cattle require 40 kg of fodder each per day that comprises forage and additional fodder. The expert suggested that, to stay healthy, the cattle require 30 kg of forage each per day. For the cows, the forage fulfilment also affects the quantity of the milk they produce.

A.2.3 Milk production procedure

Farmers' cows that have been pregnant can produce milk. The first pregnancy usually occurs after the cow's age reaches two years. The quantity of milk produced is determined by several factors (i.e., age, genetics and forage) as described in equation A.2.

$$Qm_i = \begin{cases} MaxProd_i * ProdEff(NumPreg_i) * \overline{Forage}_i, & PregPeriod < 7 \text{ month} \\ 0, & PregPeriod > 7 \text{ month} \end{cases} \quad (A.2)$$

Qm_i denotes the quantity of milk produced by cow i in a day. The $PregPeriod$ variable represents how long the given cow has been pregnant. The farmers usually stop milking a cow that has been pregnant for 7 months and restart the milking process after it gives birth. Hence, the milk production during this period is zero. $MaxProd_i$ denotes the maximum milk production and reflects the genetic attributes of the given cow. $ProdEff$ represents the percentage of milk that is produced by the given cow. It is a function of the number of pregnancies ($NumPreg_i$) as a proxy for age. The expert suggested that a cow achieved its maximum milk production after the second pregnancy ($ProdEff(2) = 100\%$) and the milk production then decreased in the subsequent pregnancies. Actual measurements to establish this function were not available but the expert suggested that it would be reasonable to assume that the milk production was decreasing linearly. We also assume that the milk production is proportional to \overline{Forage}_i , which represents the average forage fulfilment (between 0 and 1) of cow i . The average forage fulfilment of 1 means that the given cow always obtains sufficient forage throughout its lifetime.

In addition to the quantity, we also consider the milk quality in our simulation. This variable determines the milk price per litre received from the cooperative. The expert suggested that the proportion of forage in the total fodder determined the milk quality. We assume that the quality of milk increases linearly with the increase in the proportion of forage. When the proportion is 75% or higher, the quality of milk reaches its maximum (i.e. 100%).

A.2.4 Hypothesized cattle selling and buying processes

Decisions regarding how many cattle should be retained are the most important decision made by the farmer agents because it would affect the amount of forage required, cattle weight, mortality and the amount of additional fodder used (Gross, et al., 2006). In our simulation, three separate processes determine how the farmers buy or sell their cattle. In the first process, the decision to sell or buy cattle is triggered by the forage availability (Gross, et al., 2006). In the second process, this decision is triggered by the cattle's age (Rasch, Heckeley, Oomen, & Naumann, 2016). Finally, in the last process, it is triggered by the farmer's financial condition (Boone et al., 2011).

In common with earlier studies (e.g., Gross, et al. (2006)), for the first process, we assume that the forage availability is a trigger for the farmers to sell or buy cattle. When the forage is less available (e.g., during a drought), they sell some of their cattle and, conversely, buy new cattle (particularly cows) when the forage becomes more available. We assume that the farmers sell or buy their cattle to an external agent outside the system and not to other farmers in common with earlier studies (Boone, et al., 2011).

When the average forage collected is not sufficient to feed all of their cattle, they will start to sell their cattle. According to the experts, the farmers will prioritize the sale of the bulls first because they do not generate regular income. They will start to consider selling their cows only when they do not have any more bulls. When selling the cows, farmer agents compare the

potential income they can get by feeding less forage but retaining all of their cows (equation A.3) with the potential income they can get by feeding sufficient forage but selling some of their cows (equation A.4).

$$income^{retain} = n_{cow} \left(\left(\overline{Qm} * \overline{MP} * \frac{\overline{Fc}}{30} \right) - \left(\left(10 + 30 - \frac{\overline{Fc}}{n_{cow}} \right) * AfP \right) \right) \quad (A.3)$$

$$income^{sell} = \frac{\overline{Fc}}{30} ((\overline{Qm} * \overline{MP}) - (10 * AfP)) \quad (A.4)$$

n_{cow} denotes the number of cows currently owned by a farmer. \overline{Fc} represents the average forage obtained by the farmer and $\overline{Fc}/30$ represents the maximum number of cows the farmer can retain for the given forage availability. \overline{Qm} , \overline{MP} and AfP represent the average milk production per cow, the average milk price per litre and the additional fodder price, respectively. In equation A.3, the farmer has more cows to produce milk but suffers a production penalty owing to the lack of forage and must pay more for additional fodder. In equation A.4, the farmer has fewer cows, but each cow can produce more milk and the agent does not need to buy additional fodder. If $income^{sell} > income^{retain}$ then the farmer will decide to sell the cows and vice versa.

When selling cattle, we assumed that the farmers will prioritise the sale of the oldest cattle first as described in earlier studies (e.g., Boone, et al. (2011)). For the bulls, an older bull usually has more live weight (Quang, Schreinemachers, & Berger, 2014) and is more valuable. Older cows have more live weight, but older cows are usually considered less productive.

On the other hand, if the farmers can collect more forage than is needed then they start to consider buying more cows. The number of new cows a farmer is willing to buy is proportional to the additional cows that can be fed using the excess forage. For example, if the forage excess is enough to feed one more cow then the farmer agent will consider buying one additional cow. The constraints in the buying decisions are the pen capacity and the farmer's money. If the farmer agent owns sufficient pen capacity for its current cows and the new cows, then it just needs to have sufficient money to buy the cows. However, if the farmer agent does not have sufficient pen capacity then it must have sufficient money to buy the cows and to increase the pen capacity. The farmer agent's ability to increase pen capacity is also limited by the land availability on the patch where it is living. The fertility and productivity of newly bought cows are assumed to be random.

In the second process, the selling decision is based on the cattle's age. In general, the bulls are sold at two years old. According to experts, farmers believe that the bulls have reached their optimum live weight at this age. Meanwhile, the cows are culled when they reach the age of 10 years. At that age, it is believed that the milk productivity of the given cow has become too low.

In the third process, the cattle selling decision is triggered by the farmer's financial condition. Each month, the farmer agent calculates its living expenses and estimates the amount of money it will have at the end of the month. The living expense value is calculated by multiplying the

1
2
3 number of the farmer agent's family members and the standard cost of living in the area. If the
4 estimated income is less than the living expense value, then the farmer agent starts to consider
5 selling its cattle. As in the first process, the farmers are assumed to sell the bulls first. In this
6 process, we also assume that the farmers select the cattle to sell based on their age. This selling
7 process is repeated until the farmer agent's money deficit is covered.
8
9

10 In all of those three processes, we identify the maximum amount of money that can be earned
11 by farmers from selling their cattle using the survey data. The actual price received by the
12 farmer agent is assumed to be proportional to the age of cattle being sold. On the other hand,
13 the price that must be paid by the farmers to buy a new cow is assumed to be constant and the
14 value is determined using the survey data.
15
16

17 It should be noted that the buying and selling decision rules described here are hypothesized
18 (i.e. not elicited from farmers). The elicited buying and selling decision rules are described in
19 the paper.
20
21

22 *A.2.5 Cow reproduction*

23
24 Prior studies incorporate a fixed time schedule (e.g., annually) or growth rate (e.g., increase the
25 population by 10% every year) for cow reproduction (e.g., Martin, et al. (2016)). In our study,
26 we considered a heterogeneous cow fertility factor. In the cow reproduction process, the
27 farmers artificially inseminate those cows who are two years old or older and not pregnant at
28 the beginning of each simulation month. The successfulness of the artificial insemination
29 process depends on the cow's fertility, which is represented by the services per conception
30 variable. If the artificial insemination fails, then this process would be repeated in the
31 subsequent month.
32
33

34 If the artificial insemination process is successful, then the pregnancy process lasts for nine
35 months. The cow then gives birth to either a male or a female calf, each with 50% probability.
36 If the cow gives birth to a female calf, then the new-born calf inherits the milk productivity and
37 fertility of its mother.
38
39

40 *A.2.6 The retirement and succession of farmer households*

41
42 There are two main factors affecting retirement and succession of farmer household agent (i.e.,
43 age and financial condition). At the end of each simulation year, all farmer household members
44 who are older than the productive age are removed from the farmer household family member
45 list and the amount of family labour decreases. A farmer household agent can also acquire a
46 new family member with a probability of 1.2% (the average population growth in Indonesia).
47 A farmer household agent is deleted from the simulation if it does not have any family member
48 left or if it runs out of money and cattle.
49
50
51

52 Owing to the population growth, the farmland that was once located in the rural area is currently
53 surrounded by residential areas. The non-farmers tolerate the existence of a farmer household
54 who continues dairy farming because they are native to the area while the non-farmers are
55 mainly newcomers. The cooperative's database also shows that all of its members are multi-
56 generation farmer families. However, conflict with non-farmers could spark easily if a
57 newcomer tries to start dairy farming. This conflict is usually triggered by pollution caused by
58 manure production and potential water contamination. When a farmer household decides to
59
60

1
2
3 stop dairy farming, their land will usually be sold and converted into residential area settlement
4 or another business. Our simulation aims to replicate the reality in the case study area so the
5 probability of a new farmer agent entering the system is set to be equal to zero. However, a
6 sensitivity analysis can be run on this probability value if the survey data indicated that the
7 emergence of new farmers was quite possible.
8
9

10 *A.3 The cooperative agent*

11
12 The cooperative agent collects and grades milk from all farmer household agents. We assume
13 that the cooperative determines the milk-buying price as a linear function of milk quality,
14 ranging from Rp 3350 to Rp 5200 per litre (Rp is Indonesian currency). Based on the discussion
15 with the experts, the cooperative sells the milk to the milk processing industry at a fixed price.
16 The actual buying price from the milk processing industry is unknown but the experts estimated
17 that it was approximately Rp 5500 per litre. The experts agreed that the cooperative's daily
18 operational costs can be assumed to be fixed regardless the total volume of milk they handle.
19 Hence, it is more profitable if they can operate at full capacity.
20
21
22

23 **References**

- 24
25 Bahar, S. (2014). Produktivitas hijauan pakan untuk produksi sapi potong di Sulawesi Selatan.
26 *JITV*, 19(3)
27
28 Boone, R., Galvin, K., BurnSilver, S., Thornton, P., Ojima, D., & Jawson, J. (2011). Using
29 coupled simulation models to link pastoral decision making and ecosystem services.
30 *Ecology and Society*, 16(2)
31
32 Gross, J. E., McAllister, R. R. J., Abel, N., Smith, D. M. S., & Maru, Y. (2006). Australian
33 rangelands as complex adaptive systems: a conceptual model and preliminary results.
34 *Environmental Modelling & Software*, 21(9), pp. 1264-1272.
35
36 Martin, R., Linstädter, A., Frank, K., & Müller, B. (2016). Livelihood security in face of
37 drought – Assessing the vulnerability of pastoral households. *Environmental Modelling*
38 & *Software*, 75(Supplement C), pp. 414-423.
39 doi:<https://doi.org/10.1016/j.envsoft.2014.10.012> Retrieved from
40 <http://www.sciencedirect.com/science/article/pii/S1364815214003028>
41
42 Quang, D. V., Schreinemachers, P., & Berger, T. (2014). Ex-ante assessment of soil
43 conservation methods in the uplands of Vietnam: An agent-based modeling approach.
44 *Agricultural Systems*, 123, pp. 108-119.
45
46 Rasch, S., Heckelei, T., Oomen, R., & Naumann, C. (2016). Cooperation and collapse in a
47 communal livestock production SES model – A case from South Africa. *Environmental*
48 *Modelling & Software*, 75(Supplement C), pp. 402-413.
49 doi:<https://doi.org/10.1016/j.envsoft.2014.12.008> Retrieved from
50 <http://www.sciencedirect.com/science/article/pii/S136481521400365X>
51
52
53
54
55
56
57
58
59
60

Appendix B: Survey Part 1 (model parameterisation)

This section presents some of the questions that are used to validate assumptions in the base model.

- 1) How old are you now?
- 2) Is dairy cow farming the only source of income for your family? (a) Yes; (b) No
- 3) If you have other sources of incomes, please estimate the proportion of these other incomes from your total income?
- 4) Do your family members assist you in dairy farming? (a) Yes, How many people? (b) No
- 5) Did you ever hire outside workers (not your family member) to assist you in dairy farming? (a) Yes; (b) No
- 6) If yes, then how many outside workers did you usually hire?
- 7) If yes, then for how many days in a month did you hire this outside worker?
- 8) How many cows do you have? Please also mention the age of the cows that you can remember well.
- 9) How many bulls do you have? Please also mention the age of the bulls that you can remember well.
- 10) Please choose one of your cows which characteristics you can remember best. What is the highest milk production ever produced by that cow? At what age does that cow reach this level of production? How old is that cow now? How much milk does it produce currently? How many times should the artificial insemination be given to make that cow pregnant?
- 11) How do you acquire the forage for your cattle (you can give more than one answer)? (a) I grow the forage; (b) I collect the forage from the areas surrounding my village; (c) I hire other people to collect the forage; (d) I buy the forage from a trader; (e) other, please explain
- 12) If you grow the forage, then how much land do you have (your own land and/or land that you rent)?
- 13) Can you produce sufficient forage to feed all of your cattle? (a) Yes; (b) No
- 14) If your answer is no, then how much land you supposedly have to produce sufficient forage (assume that the forage growth rate is constant)?
- 15) On average how much forage do you give to all of your cattle in one day?
- 16) Do you also feed your cattle with additional fodder? (a) Yes; (b) No.
- 17) If your answer is yes then how do you acquire the additional fodder for your cattle (you can give more than one response)? (a) I produce it by myself; (b) I buy it from supplier; (c) other, please explain
- 18) If you buy the additional fodder from the supplier, what is the price per kilogram of additional fodder?
- 19) On average how much additional fodder do you give to your cattle in one day?
- 20) If you have sold a cow in the past year, then what is the highest selling price you receive? How old was your cow when it is sold?
- 21) If you have sold bull in the past year, then what is the highest selling price you receive? How old was your bull when it is sold?
- 22) If you have bought heifer in the past year, then what is the highest buying price you pay?
- 23) From your experience, what is the highest milk price you received in the past year?
- 24) From your experience, what is the lowest milk price you received in the past year?

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

25) The followings are some factors that might force you to quit dairy farming. Please state how big the likelihood of these factors to force you to quit from dairy farming (5 very likely, 1 very unlikely)

a) My children do not want to continue my dairy farming business	1	2	3	4	5
b) The income from dairy farming is not sufficient for daily living	1	2	3	4	5
c) It is become more difficult to obtain sufficient forage	1	2	3	4	5
d) Competition with new farmers become more intense	1	2	3	4	5
e) The cattle quality (genetic breed) is decreasing	1	2	3	4	5

For Peer Review Only

Appendix C: Detailed statistical analyses

Equation 1

Variables Entered/Removed^b

Mode	Variables Entered	Variables Removed	Method
1	Milk Price Per Liter (IDR/Liter), Sqrt Additional Forage Obtained (Kg-2) ^a		Enter

a. All requested variables entered.

b. Dependent Variable: Additional Cow Wanted (Head)

Model Summary

Mode	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.384 ^a	.147	.136	3.75653

a. Predictors: (Constant), Milk Price Per Liter (IDR/Liter), Sqrt Additional Forage Obtained (Kg-2)

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	361.219	2	180.609	12.799	.000 ^a
	Residual	2088.503	148	14.112		
	Total	2449.722	150			

a. Predictors: (Constant), Milk Price Per Liter (IDR/Liter), Sqrt Additional Forage Obtained (Kg-2)

b. Dependent Variable: Additional Cow Wanted (Head)

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1.603	4.889		-.328	.743
	Sqrt Additional Forage Obtained (Kg-2)	.095	.019	.376	4.953	.000
	Milk Price Per Liter (IDR/Liter)	.001	.001	.084	1.108	.270

a. Dependent Variable: Additional Cow Wanted (Head)

Equation 2

Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	612.398			
Final	82.996	529.402	3	.000

Pseudo R-Square

Cox and Snell	.554
Nagelkerke	.750
McFadden	.601

Likelihood Ratio Tests

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	249.166	166.171	1	.000
Prob_Die	612.353	529.357	1	.000
Drought_Length	83.005	.010	1	.922
Forage_Sufficiency	83.021	.025	1	.873

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

Parameter Estimates

Decision ^a	B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp(B)	
							Lower Bound	Upper Bound
Sell Intercept	-6.342	.646	96.408	1	.000			
Prob_Die	11.442	.897	162.722	1	.000	9.317E4	16061.248	540526.900
Drought_Length	.001	.006	.010	1	.922	1.001	.988	1.013
Forage_Sufficiency	.106	.666	.025	1	.873	1.112	.301	4.106

a. The reference category is: Retain.

Classification

Observed	Predicted			Percent Correct
	Sell	Retain		
Sell	331	63		84.0%
Retain	17	244		93.5%
Overall Percentage	53.1%	46.9%		87.8%

Equation 3

Variables Entered/Removed^b

Mode	Variables Entered	Variables Removed	Method
1	High_Fertility, Not_Pregnant, Young_Cow ^a		Enter

a. Tolerance = .000 limits reached.

b. Dependent Variable: Priority

Model Summary

Mode	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.989 ^a	.978	.978	.33753

a. Predictors: (Constant), High_Fertility, Not_Pregnant, Young_Cow

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2630.125	3	876.708	7.695E3	.000 ^a
	Residual	57.875	508	.114		
	Total	2688.000	511			

a. Predictors: (Constant), High_Fertility, Not_Pregnant, Young_Cow

b. Dependent Variable: Priority

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.766	.030		92.701	.000
	Young_Cow	4.000	.030	.873	134.076	.000
	Not_Pregnant	-1.750	.030	-.382	-58.658	.000
	High_Fertility	1.219	.030	.266	40.851	.000

a. Dependent Variable: Priority

Excluded Variables^b

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	Old_Cow	∞000
	Pregnant	∞000
	Low_Fertility	∞000

a. Predictors in the Model: (Constant), High_Fertility, Not_Pregnant, Young_Cow

b. Dependent Variable: Priority

Peer Review Only