2014

A framed field experiment about policy measures – Testing the effectiveness of rewards or punishments with different probabilities as incentives in palm oil production

Stefan Moser and Oliver Mußhoff

EFForTS discussion paper series Nr. 5

Moser, Stefan; Mußhoff, Oliver: A framed field experiment about policy measures : testing the effectiveness of rewards or punishments with different probabilities as incentives in palm oil production
Göttingen : GOEDOC, Dokumenten- und Publikationsserver der Georg-August-Universität, 2014
(EFForTS discussion paper series 5)

Verfügbar:
PURL: http://resolver.sub.uni-goettingen.de/purl/?webdoc-3937

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

Erschienen in der Reihe
EFForTS discussion paper series

ISSN: 2197-6244

Herausgeber der Reihe
SFB 990 EFForTS, Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems (Sumatra, Indonesien) - Ökologische und sozioökonomische Funktionen tropischer Tieflandregenwald-Transformationssysteme (Sumatra, Indonesien)

Georg-August-Universität Göttingen
Johann-Friedrich-Blumenbach Institut für Zoologie und Anthropologie, Fakultät für Biologie und Psychologie

Abstract: Palm oil production creates negative externalities, e.g., through intensive fertiliser application. If policy wants to limit externalities, an effective, sustainable and efficient measure seems desirable. Embedded in a framed field experiment in Indonesia, we apply a business simulation game to test ex ante several incentives for reducing the use of fertiliser in palm oil production. These incentives are arranged in the form of different designs, i.e., either a reward or punishment, varying in their magnitude and probability of occurrence but constant in the effect on expected income. Results show that participants react significantly different depending on the incentive design. A high reward with a low probability to occur has been found to be the most effective and sustainable incentive design. For efficiency, a low and certain reward is indicated to be the best design.

Keywords: Policy influence analysis, effective incentive, framed field experiment, business simulation game, palm oil production, Indonesia

A framed field experiment about policy measures: Testing the effectiveness of rewards or punishments with different probabilities as incentives in palm oil production

Stefan Moser and Oliver Mußhoff

EFForts Discussion Paper Series

No. 5 (June 2014)

Funded by the German Research Foundation (DFG) through the CRC 990 "EFForTS, Ecological and Socioeconomic Functions of Tropical Lowland Rainforest Transformation Systems (Sumatra, Indonesia)"

www.uni-goettingen.de/en/310995.html

SFB 990, University of Goettingen

Berliner Straße 28, D-37073 Goettingen, Germany

ISSN: 2197-6244

Managing editors:

At the University of Goettingen, Germany

Prof. Dr. Christoph Dittrich, Institute of Geography, Dept. of Human Geography

(Email: christoph.dittrich@geo.uni-goettingen.de)

Dr. Stefan Schwarze, Dept. of Agricultural Economics and Rural Development

(Email: sschwar1@gwdg.de)

At the Universities of Bogor and Jambi, Indonesia

Prof. Dr. Zulkifli Alamsyah, Dept. of Agricultural Economics, Faculty of Agriculture, University of Jambi

(Email: zalamsyah@unja.ac.id)

Dr. Satyawan, Sunito, Dept. of Communication and Community Development Sciences, Faculty of Human Ecology, Bogor Agricultural University (IPB)

(Email: awansunito@gmail.com)

Table of contents

List of Figures

List of Tables

A framed field experiment about policy measures: Testing the effectiveness of rewards or punishments with different probabilities as incentives in palm oil production

Stefan Moser, Oliver Mußhoff[1]

Abstract

Palm oil production creates negative externalities, e.g., through intensive fertiliser application. If policy wants to limit externalities, an effective, sustainable and efficient measure seems desirable. Embedded in a framed field experiment in Indonesia, we apply a business simulation game to test ex ante several incentives for reducing the use of fertiliser in palm oil production. These incentives are arranged in the form of different designs, i.e., either a reward or punishment, varying in their magnitude and probability of occurrence but constant in the effect on expected income. Results show that participants react significantly different depending on the incentive design. A high reward with a low probability to occur has been found to be the most effective and sustainable incentive design. For efficiency, a low and certain reward is indicated to be the best design.

Keywords: Policy influence analysis, effective incentive, framed field experiment, business simulation game, palm oil production, Indonesia

[1] Contact: Stefan Moser, University of Göttingen, Department of Agricultural Economics and Rural Development, Platz der Göttinger Sieben 5, 37073 Göttingen, Phone +49 (0)551 39-4844, email: smoser@gwdg.de
Oliver Mußhoff, University of Göttingen, Department of Agricultural Economics and Rural Development, Platz der Göttinger Sieben 5, 37073 Göttingen, Phone +49 (0)551 39- 4842, email: oliver.musshoff@agr.uni-goettingen.de

1. Introduction

Palm oil is used widespread, for example in food products, cosmetics or as biofuel. With a production volume of approximately 50 million tonnes in 2012, it is the most significant vegetable oil in the world (FAOSTAT, 2013). Starting in the 1960s, with a worldwide production of several million tonnes, palm oil production has grown exponentially, doubling every 10 years. With 23.6 million tonnes in 2012, Indonesia is the largest palm oil producer in the world, working towards increasing its production by as much as 40 million tonnes by 2020 (UNCTAD, 2013).

However, there are negative externalities caused by palm oil production. Koh and Wilcove (2008) state that the expansion of palm oil production leads to deforestations, as well as to significant losses in biodiversity. Fitzherbert et al. (2008) additionally describe an effect on habitat fragmentation and pollution including greenhouse gases. Reijnders and Huijbregts (2008) found significant CO_2 emissions caused by palm oil production, along with other factors caused by intensively applying fertiliser. The exaggerated fertiliser use leads to further negative externalities in palm oil production. For example, Sekhon (1995) shows that developing countries with an increased usage of N-fertiliser have more problems with groundwater pollution. Furthermore, the use of fertiliser in humid, tropical climates can increase NOx emissions, which are major contributors to global warming (Veldkamp and Keller, 1997; Keller and Matson, 1994; Veldkamp et al., 2008; Palm et al., 2002). Moreover, this fertiliser can cause ground-level ozone in tropical oil palm plantations, high concentrations of which can be detrimental to human health (Hewitt et al., 2009). For limiting these externalities from the extensive use of fertiliser in palm oil production, an effective, sustainable and efficient incentive system seems desirable. In this paper, we use the terms effectiveness of an incentive to refer to the strength of participants' reaction to an inventive, whereas sustainability refers to the persistency of this reaction. Moreover, efficiency of an incentive refers to the summarised costs and benefits for all affected stakeholders (Balliet et al., 2011).

Incentives differ in several ways, making them different in their effectiveness, sustainability and efficiency. One differentiation is if there is a reward for desired behaviour or a punishment for undesired behaviour. Another differentiation is the probability of occurrence when behaviour is desired or undesired, respectively. Incentives also differ in their magnitude. Balliet et al. (2011) lists further possible differentiations for incentives designs, e.g., costs for giving incentives, centralized versus decentralized source of incentives, matching procedures, iterations, type of dilemma and cost-to-fine ratio. Sutter et al. (2010) find that endogenously chosen incentive mechanisms increase cooperation more than exogenously chosen ones. A paper from Herrmann et al. (2008) shows that the cultural context also matters for the incentive efficiency. All of these differences and their interaction effects may influence the overall effectiveness and efficiency of incentives.

The objective of this paper is to analyse ex ante the effectiveness, sustainability and efficiency of different incentive designs for farmers, of which little is known. More specifically, we investigate in the use of incentives for fertiliser reduction for the case of small-scale palm oil producers on Indonesia's island of Sumatra. The effectiveness, sustainability and efficiency of reward and punishment with respect to the influence of magnitude versus the probability of occurrence for incentives are analysed. An ex ante analysis is a low-cost method which allows for testing an incentive before it is applied. To do so, we apply a business simulation game within a framed field experiment.

This paper contributes to the current literature in several ways. First, we are testing reward and punishment with several fixed probabilities of occurrence in one framed field experiment. This is enabling for the direct comparison of the effectiveness and efficiency of these particular designed incentives. Furthermore, a framed field experiment allows for a higher generality of the results compared to a standard laboratory experiment (Harrison and List, 2004). The second novelty is that policy measures are tested experimentally in Indonesia. This enables a low-cost ex ante evaluation of certain measures within the cultural context of Indonesia. Third, these novelties demand a methodological approach for achieving a realistic decision situation in combination with a clear causality of the measured effects. Therefore, a multi-period business simulation game is developed and applied for analysing the named incentives. For meaningful conclusions, the resulting panel data needs to be analysed with a matching procedure and a fixed effect regression.

The remainder of the paper is structured as follows: In section 2, the literature is discussed and hypotheses are generated. In section 3, the method and experimental design is illustrated. Section 4 gives a description of the context and the sample recruitment, while section 5 states the data analysis. Section 6 presents the results along with discussion. Finally, section 7 closes with a conclusion and an outlook.

2. Literature and hypothesis generation

To internalise the negative externalities through palm oil production an appropriate incentive system is requested. Among other things, the design of such an incentive can differ in its probability of occurrence, its magnitude or in being a reward or a punishment. Problems with enforcing laws concerning probabilities and magnitudes of sanctions have been discussed theoretically by Polinsky and Shavell (1999). In combination, magnitude and probability of occurrence give the effect on the expected income, i.e., participants average cost or reward for cooperative or uncooperative behaviour, respectively. In theory, as long as the effects on expected income are held constant, a rational, risk neutral agent will behave independently of the probability and the magnitude of an incentive (Becker, 1974). Following this argumentation and assuming a rational, risk neutral agent, it makes no difference on the behaviour if the incentive is designed as reward or as punishment, and also it makes no difference if the magnitude of an incentive is increased while its probability of occurrence is decreased, as long as the effects on expected income are held constant. In contrast to that, the literature of experimental economics shows that the reactions towards an incentive do not only depend on the effects on expected income. The reactions also depend on incentive design, i.e. rewarding or punishing, the probability of occurrence and how the magnitude of an incentive is sized (Balliet et al., 2011). Subsequently, we discuss the literature about rewarding or punishing and about magnitude and probability of occurrence.

For the consequences of rewarding or punishing, extensive literature exists. According to Heath et al. (1999) people are more motivated to avoid losses than to gain profits, with the precondition that there is a reference point which determines when a result is a loss. Kahneman (2011) gives examples for loss aversion, e.g., for cab drivers in New York (Camerer et al., 1997), for Swiss messenger services (Fehr and Goette, 2007) or for professional golf players (Pope and Schweitzer, 2011). According to these papers, punishing should be more effective than rewarding. On the contrary, a meta-analysis involving 187 effect sizes from Balliet et al. (2011) shows reward and punishment having similar positive effects on cooperation, even though differences occur depending on the context. A framed field experiment from Ibanez and Martinsson (2008), which combines reward and punishment, shows that coca farmers react more strongly to increasing relative profits than to an increasing probability of occurrence for punishment. To summarise, there is evidence that the effectiveness of reward and punishment differs in certain contexts, but they do not differ in general.

According to Kahneman (2011), at a constant magnitude a low loss probability leads to risk avoiding behaviour, whereas a high loss probability causes risk loving behaviour. Anderson and Stafford (2003) found that for the same expected punishment, raising the magnitude causes more reaction than raising the probability of occurrence. Furthermore, Block and Vernon (1995) found that for punishment, the behaviour depends on the participating group, but usually the reaction to the magnitude is stronger than to its probability of occurrence. The literature regarding the effectiveness of rewards with low probability of occurrence is not that extensive. Kahneman (2011) claims risk loving behaviour for rewards with low probability, whereas there is risk avoiding behaviour for rewards with high probability at a constant magnitude. Volpp et al. (2008) gave incentives for reducing weight for overweight individuals. They found people reacting significantly to the incentive, but not significantly different between reward with a high and uncertain magnitude or a low and certain reward. In summary, we expect a stronger reaction to the magnitude than to the probability of occurrence for punishment as well as for rewarding. This leads to our first hypothesis:

H1: For the same effect on expected income, the effectiveness of an incentive is independent of its design, i.e., reward or punishment, with different magnitudes and probabilities of occurrence.

In their meta-analysis, Balliet et al. (2011) discuss the sustainability of incentives, i.e. the persistency of an incentive effectiveness, on cooperation. They found incentives being more effective for iterated experiments compared to one-shot experiments. They explain this effect on both individual learning and group learning processes. They do not indicate whether reward or punishment is more effective in terms of sustainability. Moreover, they do not discuss how the magnitude or the probabilities of incentives affect its sustainability. To find out more about the relation of reward, punishment and sustainability, we formulate the second hypothesis as follows:

H2: For the same effect on expected income, the sustainability of an incentive is independent of its design.

The effect of incentives on efficiency is another dimension in this analysis. In the literature, a public good game is often applied to find out more about the efficiencies of incentives. In their meta-analysis Balliet et al. (2011) summarise that punishment is efficient, at least in the long run, but they do not directly compare the efficiency of reward and punishment. Rand et al. (2009) and Sefton et al. (2007) show that reward is more efficient than punishment. However, little is known as to how efficiency develops if the magnitude of the incentives is raised, and the probability of occurrence of these incentives is decreased while the effects on the expected income are maintained constantly. This leads to our third hypothesis:

H3: For the same effect on expected income, the efficiency of an incentive is independent of its design.

3. Method and experimental design

To test for these hypotheses, a framed field experiment is conducted. Harrison and List (2004) define such experiments as conventional lab experiments, but with a nonstandard subject pool as well as with field context. Roe and Just (2009) discuss the importance of high internal and external validity for economic research. Internal validity occurs if a researcher can argue that observed correlations are causal. In the present study, this is achieved by ensuring a constant experimental structure and only changing the incentives design. External validity refers to the ability to generalize the relationships found in a study to other persons, times and settings. According to Levitt and List (2007) adapting the experiment to a realistic decision context creates generalisability. A business simulation game allows for describing such a relatively realistic decision situation in combination with a clear causality of the measured effects.

Configuration of the applied business simulation game is a naturally occurring trade-off between being realistic on the one hand and being workable with participants and various circumstances on the other hand. The necessary assumptions for meeting these conditions are explained to the participants. To avoid confusion among participants, it is stressed that some of these assumptions are not completely realistic. The resulting structure of the business simulation game explained to the farmers is as follows: It is simulated that each farmer manages a one hectare palm oil plantation that is already established and producing. The farmer has to do so for 10 rounds, where each round is an equivalent for one year. At the beginning of each round, the participant has to decide how much fertiliser to use. It is assumed that the yield on the simulated plantation depends solely on the amount of fertiliser used in the corresponding round, i.e., weather-related yield fluctuations, emerging diseases, etc. are not taken into consideration. Each participant gets the same yield for the same amount of fertiliser. To generate profit, the entire harvest is sold at the end of the round. At the beginning of the experiment, the participants were told that there is a policy measure in the form of a subsidy to raise farmers' income and that this policy measure can change during the experiment. If necessary, liquidity is always available and there is a zero interest rate for the gained profit as well as for credits. Since the change in policy measure always occurs after the fifth round, rounds 1-5 and rounds 6-10 are named the first and second sequence, respectively. For causing a realistic behaviour, participants were given an incentive. Thus, the gained profits from the ten rounds are added up and 5 per cent of these profits is paid to them in the form of a shopping voucher.

The proceeding for each round is constant. The first step is the participants' decision of how much fertiliser to use. In the second step, a random process fixes the price for the yield for the corresponding round. In the third step, an enumerator determines the achieved yield and calculates the profit. A more detailed explanation for these steps follows.

The first step in the business simulation game is the participants' decision of how much fertiliser to use in the current round. This is the only decision a participant has to make in the current round. Except for the first round, farmers do not yet know the product price of the current round when they make this decision. The total amount of fertiliser used solely determines the yield on the simulated oil palm plantation. A quadratic production function is applied to calculate the yield. The functional form is $f(x) = 10 + 0.05x - 0.000055x^2$, where $f(x)$ and x represents the yield amount and fertiliser respectively. To reduce the mental effort for the participants, only 10 kg multiples for fertiliser were allowed to be used. The corresponding yields are transferred into a table and a graph (Figure A1), which is handed out to the farmers at the beginning of the experiment. The formula behind this table was not shown and explained to the farmers. The table starts with 0 kg and ends with 590 kg, even though participants were allowed to use amounts of fertiliser beyond this scale. Anyway, the maximum yield is achieved with 440 kg of fertilizer. A report by the FAO (2005) recommends similar amount of fertiliser being used for palm oil plantations. In our business simulation game the possible yields range from 10 tons to a maximum of 21.4 tons palm oil bunches, which corresponds to the approximate annual yield per hectare found in the literature (Fairhurst and McLaughlin, 2009; Jelsma et al., 2009).

The second step in the experiment is the evaluation of the price. Only in the first round, the price of IDR 1,800 per tonne palm oil bunch is told to the participants before their decision about the amount of fertiliser used, whereas for all the other rounds, the price was evaluated after the farmer decided for their use of fertiliser. As a basis for this, a price of IDR 1,710 per kg of palm oil bunch in the year 2008 (Jelsma et al., 2009) was found in the literature. To make the numbers in the experiment better manageable, we decided to replace the yield price per kg with the yield price per tonne. For all of the following rounds, the price rises or falls by IDR 200 per ton of palm oil bunch based on the price of the previous round. This means, e.g., for the second sequence, that the price either rises to IDR 2,000 or falls to IDR 1,600 with a 50% probability, respectively. Thus, the price development follows an arithmetic Brownian Motion (Poitras, 1998) until the 10^{th} round, with a minimum of IDR 0 and a maximum of IDR 3,600. To evaluate whether the price rises or falls, participants have to draw a ball from a bag which contains three green and three red balls. Prices rise for drawing a green ball and fall for drawing a red ball. To achieve a distribution of the price developments, these prices were determined separately for each participant.

The third step of calculating the realized profit for the round is straightforward. The revenue in a particular round equals the achieved yield multiplied by the product price in the corresponding round. Since palm oil bunches have to be processed promptly after harvesting, there is no option for storing and thus, the whole production has to be sold in the corresponding round. The cost for fertiliser is fixed at IDR 10 per kg. The revenue and costs for fertiliser are then summed. Additionally, there is the possibility of having a policy measure that also contributes to the profit. This measure will be explained in the next chapter. After the profit is calculated, the next round starts and the participant must again decide how much fertiliser to use. After 10 rounds, the experiment is over.

The policy measure can differ between the first and the second sequence or between participants. For the first sequence, the participants always receive a fixed subsidy of IDR 10,000 per round, without regard for their use of fertiliser. For the second sequence, this public subsidy can change to one out of four policy measure for reducing the use of fertiliser; during the second sequence, this measure remains constant. At the beginning of the first round, the farmers are informed that subsidies may change once during the experiment, but are not given any further information. At the beginning of the sixth round, the policy measure is explained to the farmers. The treatments for the second sequence are structured as follows:

Control Treatment: The public subsidy of IDR 10,000 proceeds without any restriction, there is no change.

Treatment A: The subsidy is replaced by a compensation payment of IDR 10,000. It is paid if farmers use 120 kg of fertiliser or less in the corresponding round.

Treatment B: The subsidy proceeds. Additionally, there is a measure that punishes farmers who uses 130 kg of fertiliser or more in the corresponding round with IDR 10,000. This equals a deduction of the received subsidy.

Treatment C: The subsidy is replaced by a compensation payment of IDR 100,000. It is paid if farmers use 120 kg of fertiliser or less in the corresponding round, but with an occurrence probability of only 10%.

Treatment D: The subsidy proceeds. Additionally, there is a measure that punishes farmers who uses 130 kg of fertiliser or more in the corresponding round with IDR 100,000, but with an occurrence probability of only 10%.

The trigger of 120 kg, as well as the magnitudes of IDR 10,000 or IDR 100,000, is chosen arbitrarily at a level where considerable reaction is expected. It is remarkable that the effect on expected income is equal for all treatments, except for the control treatment. This means that a rational and risk neutral participant would behave independently of the treatment (Polinsky and Shavell, 1999). The idea of the policy measure standing behind treatment A and treatment B is certain rewarding for desired or certain punishing for undesired behaviour. For treatment C, the idea is close to a outcome-based incentive (Tute, 2005), where the probability of occurrence reflects the possibility of failing defined outcomes, even when serious action is taken to achieve them. Treatment D considers the situation when controlling is imperfect and only successful with a 10% chance. For treatments C and D, the incentives are equipped with a 10% probability to occur. To determine if the participant is imposed or not, a similar procedure as for the price determination was applied. Participants have to draw from a bag with nine blue and one yellow ball. If they draw the yellow ball, the incentive occurs, whereas nothing happens if a blue ball is drawn.

4. Context and sample recruitment

The experiment is executed in the Jambi Province on Sumatra, Indonesia. Jambi has about three million inhabitants and has an area of approximately 50,000 km². The research area is in four regencies of the Jambi Province, i.e., Tebo, Bungo, Batang Hari and Muaro Jambi. Over the last several decades, there has been a strong transformation of the landscape towards palm oil plantations (Laumonier et al., 2010; Wilcove and Koh, 2010), making it a valuable research area for the topic of this paper.

The data was collected from 29 randomly chosen villages in the research area from October to December 2012. For each village, an entire list of farmers was created. Depending on the size of the village in terms of inhabitants, between 10 and 18 small-scale farmers are chosen randomly and invited to participate in the experiment. The experiment started in the afternoon or in the evening after prayer time to accommodate local customs. The experiment was completed in available public rooms or sometimes in the house of the village head. The business simulation game is done simultaneously for the whole group. Before the experiment started, participants were asked about their socioeconomic data. To avoid conflicts among participants, every participant from each village received the same treatment. At the beginning, each participant received a questionnaire where his or her fertiliser decisions, prices and profits are filled in at appropriate time. Then an enumerator explained the procedure with the support of visual posters and the participants had the chance to ask questions. During the experiment, the participants were divided into subgroups, so each enumerator was monitoring between 3 and 5 participants. This structure enabled the participants to ask questions on a more personal level. This is especially important for cultural reasons, since participants often hesitate asking when in a large group. Closer information about how the experiment was conducted is provided in Appendix A.

In total, 328 small-scale farmers participated in the experiment. Six uncompleted questionnaires were dismissed, resulting in 322 participants for the analysis. On average, participants made a profit of IDR 37,940 per round, resulting in IDR 18,970 worth of shopping vouchers for the entirety of the business simulation game. Considering that the average daily wage for a worker is around IDR 50,000 in the research area, this seems to be a sufficient compensation for participating in this one hour experiment.

Table 1 gives the socioeconomic data of the participants. For few participants with completed questionnaires we lack in some socioeconomic data. A Kruskal-Wallis-test shows no significant differences of the participants of the five treatments in terms of sex, age, years of education and household size. This indicates a proper sample selection.

Table 1: Average socioeconomic data and fertiliser use[a)]

	Sex[d)]	Age	Education years	Household size	Fertiliser: Round 1-5	Fertiliser: Round 6-10	Participants
Control Treatment	0.81	44.16 (1.49)	7.61 (0.42)	4.49 (0.18)	278 (8.18)	288 (7.88)	71
Treatment A	0.86	44.52 (1.35)	7.60 (0.36)	4.61 (0.24)	288 (8.39)	180 (8.07)	66
Treatment B	0.83	40.85 (1.41)	8.62 (0.45)	4.19 (0.16)	275 (9.63)	169 (7.80)	60
Treatment C	0.93	42.05 (1.61)	7.44 (0.39)	4.39 (0.21)	254 (9.21)	171 (7.56)	61
Treatment D	0.94	44.43 (1.38)	7.67 (0.39)	4.44 (0.20)	244 (9.08)	161 (6.02)	64
Observations[b)]	319	319	319	313	1,610	1,610	322
Kruskal-Wallis-test[c)]	0.649	0.112	0.406	0.648	0.002	0.000	

Source: Author's data,
a) Average amount per treatment, Standard deviation in parentheses
b) N varies between the variables
c) p-values
d) 1=male, 0=female

In the first sequence of the business simulation game, participants use 268 kg of fertiliser on average. For the second sequence, the control treatment uses 288 kg, whereas the four treatment groups use 171 kg of fertiliser on average. This indicates that participants react to treatments taking place after the fifth round. Since the treatment for the first sequence is equal for each participant, it was expected that groups do not differ in their use of fertiliser in the first sequence. A Kruskal-Wallis-test shows that at the 1% level this expectation does not hold. Since the socioeconomic data are equally distributed, these differences in the use of fertiliser cannot be explained by socioeconomic data.

5. Data analysis

This section discusses the method in which the collected data is analysed. To begin with, it is shown that a matching procedure is necessary for a meaningful analysis, followed by a description of this procedure. Afterwards, the generating of the variables and the estimation of the incentives effectiveness, sustainability and efficiency are described.

In this paper, the first sequence is the baseline for estimating the incentives effectiveness, sustainability and efficiency. To ensure comparability this baseline has to be equal independent of the incentive design, i.e., treatment. Therefore, the treatments differences in fertiliser use in the first sequence, shown in Table 1, have consequences for the analysis. For example, in the fifth round, participants with treatment D have an average fertiliser use of about 250 kg, whereas participants with treatment A have an average use of 380 kg fertiliser. The trigger for the incentive in the second sequence is always 120 kg fertiliser. Therefore, treatment D participants have to reduce their use by 130 kg, whereas treatment A participants have to reduce their use by 260 kg to pass the trigger level of 120 kg of fertiliser. For adapting, treatment D participants' behaviour needs to change considerably less than treatment B participants' behaviour. Since we investigate the changes of behaviour through different designed incentives, and the baseline for our analysis is the first sequence, we cannot compare the incentives with this difference of fertiliser use in the first sequence. A Kruskal-Wallis-test for the prices shows no differences at a 5% level between the treatments in the first sequence, indicating that prices are not a source for the differences in the use of fertiliser.

5.1. Matching procedure

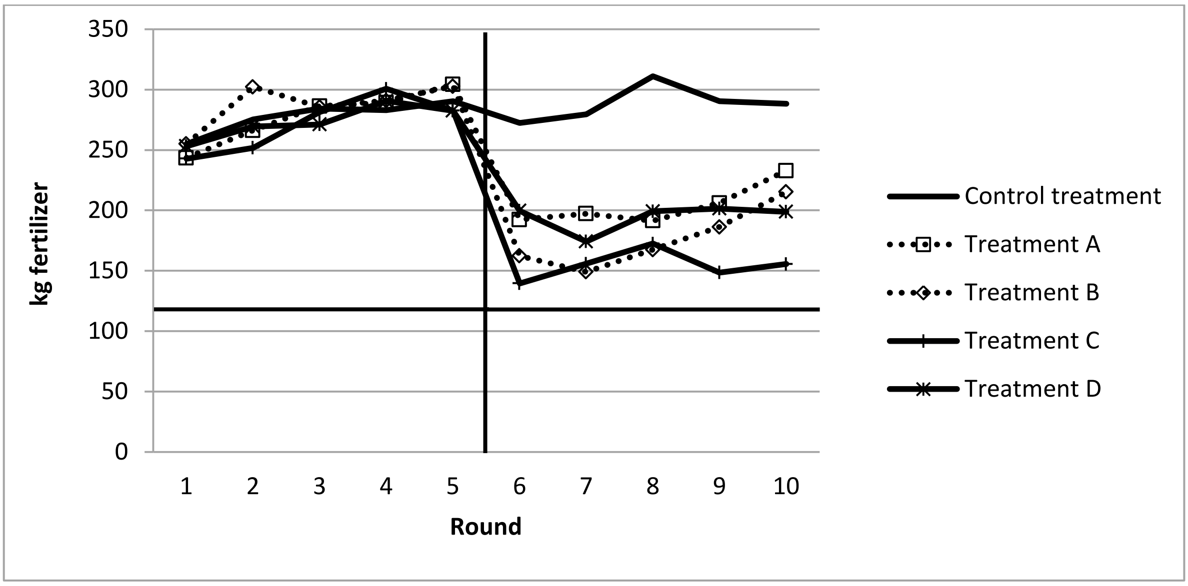
To overcome the problem of different fertiliser use in the first sequence, a minimum Euclidian distance matching is applied. Therefore, the amounts of fertiliser used in the first five rounds are matched; this approach is also used by other researchers (Tiedemann and Latacz-Lohmann, 2013). The idea behind this method is intuitive: If participants of the treatments behave similarly for rounds one through five, and then they behave differently for rounds six through ten, this difference is caused by the difference in the treatment. Treatments A, B, C and D are matched separately with the control treatment. The Euclidian distance for the matching procedure is calculated as follows:

$$d_{ij} = \sum_{k=1}^{K} \left(x_{ik} - x_{jk}\right)^2 \tag{1}$$

In Formula (1) d_{ij} is the Euclidian distance between participant i from the control treatment and participant j from the respective treatment group. k represents the round and x is the amount of fertiliser used. For the matching of participant i, the participant j with the minimum Euclidian distance is taken; this equals a nearest neighbour matching. If taken for the matching, individual j is replaced, but only at a maximum of two times. Therefore, the data of particular participants is not used too often for analysis. A more frequent use of one individual data set would automatically exclude the consideration of the data set of other participants, leading to higher loss in data. If there are two participants with the same Euclidian distance, both of them are taken. This leads to 71, 72, 75, 79 and 83 observations used for data analysis for the control treatment, treatment A, B, C and D respectively. As expected, after the matching a Kruskal-Wallis-test indicates no significant difference in the use of fertiliser at a p-value of 5 % between the treatments in the first sequence. Thus, after the matching we can use the first sequence as baseline for the analysis without hesitation. The resulting average fertiliser used per treatment is displayed in Figure 1.

Figure 1 shows the mean amount of fertiliser used per round and per treatment after the matching procedure. The horizontal line presents the 120 kg fertiliser trigger for the incentives, whereas the vertical line stands for the policy change after round five. A look at Figure 1 indicates differences between the treatments for the second sequence.

Figure 1: Average amount of fertiliser used per round and treatment after matching



Source: Author's illustration

5.2. Generating variables and regression analysis

The matched data are the basis for evaluating the differences in the treatments. To test for our hypothesis, we compare sequence one, where each participant is treated equally, with sequence two, where the different treatments take place. To analyse the effectiveness of the incentives, several variables are used. For testing the different levels of fertiliser use due to policy intervention, shift variables are introduced. These are dummy variables, one for each treatment, showing the differences between the first and the second sequence. Figure 1 indicates that the average amount of fertiliser used rises within the first sequence. Most of the participants are below the profit maximum level of around 400 kg fertiliser use. Assuming that maximising the profit influences the farmers' decisions, this rising use of fertiliser in the first sequence can be explained reasonably. This development is interpreted as a learning effect, which appears for all treatments in the first sequence simultaneously. To account for that, a learning variable is introduced for all treatments. For the first sequence, this variable starts with zero and up until the fifth round, the variable increases by one per round. Afterwards, this variable remains constant, since gained experience is not lost.

For the second sequence, the situation for this learning effect is different. For treatments A, B, C and D, the 120 kg trigger for the incentives restricts the increasing use of fertiliser. For these treatments, rounds six through 10 show whether fertiliser use remains at the decreased level caused by the incentive. If so, this development is seen as an indicator for sustainability. For the control treatment, no incentive restricts the use of fertiliser. The development influenced by the profit maximum could go on, which would be interpreted as on-going learning effect. The sustainable or learning variable is generated separately for each treatment respectively. It starts with one in round six and then increases by one for each following round.

For analysing the efficiency of the treatments, we calculate the profit difference.[2] In the second sequence, the profit difference represents the difference in profits between the first local profit maximum, which always is at the incentives trigger of 120 kg fertiliser use, and the second local profit maximum, which is around 400 kg, i.e., when participants ignore the incentive. If we assume that participants' decisions are influenced by maximising their profit, they tend towards the first local profit maximum when the profit difference is low or even negative, whereas they tend towards the second local profit maximum when the profit difference is high. For analysing efficiency, it is interesting if and how strong participants adapt their use of fertiliser depending on this profit difference. For example, if participants show a strong adaptation to this difference, a high profit difference leads to a significant increase in the participants' use of fertiliser, while vice versa a low profit or negative difference results in a decrease in fertiliser used. In doing so, they avoid much negative externalities in a situation where they lose relatively little profit. In other words, they avoid externalities when the loss in profits is low, which is quite efficient. If they would not adapt to the profit difference, the average price for avoiding negative externalities would be higher. Therefore, the strength of the adoption to the profit difference is our efficiency measurement. The profit difference is calculated separately for each treatment in the second sequence. For all treatments in the first sequence and for the control treatment in the second sequence, it is zero since no incentives occur. Efficiency effects are estimated separately for each treatment.

The profit difference is mainly driven by the product price. Therefore, estimating a price effect for each treatment would lead to multi-collinearity. For the control treatment no profit difference exists, therefore, the influence of price can be taken into account. For the first sequence, where no profit difference occurs, the price effect is estimated for all treatments simultaneously. The average profit difference is negative for each treatment. If the estimated coefficient for the profit difference has a significant size, this variable would explain a part of the shift in fertiliser use caused through the incentives. This means that our measurement in effectiveness would be partly explained through the profit difference. This would lead to a bias for the shifts of the treatment; to avoid this bias, the profit difference is corrected by diminishing the individual values by the mean value of the corresponding treatment. This measure corrects the coefficient of the shift variables for the described bias and has no effect on the other results.

[2] The efficiency of an incentive is defined as its effect on the sum of the stakeholders' gains and losses. For our experiment, this includes the direct costs for the incentives payment, the farmers' profits and the negative externalities caused through the use of fertiliser. The payments for the incentive are a transfer from a public institution to a farmer. If we assume no transaction costs, there are no efficiency losses through such transactions and we therefore ignore these payments for the efficiency analysis at hand. The farmers' profits are determined from the business simulation game. The generated negative externalities are assumed to be constant for a certain amount of fertiliser used. Since we can ignore the direct costs for the incentives, efficiency is a trade-off between farmers' profits and negative externalities through fertiliser use. To be efficient, negative externalities should be avoided when the costs in terms of farmers' profits are low. In the paper at hand, we analyse our incentives in their relative efficiency. With the data at hand, we cannot say anything about absolute efficiency, i.e. if we internalise externalities adequatly. It is possible that participants even overreact through the incentive in terms of efficiency, but we assume that this is not the case here.

A Hausman-test indicates that a random effect model does not fit. Thus, with the availability of panel data, a fixed effect regression allows us to account for unobserved, unchanging participant-specific effects:

$$y_{it} = a_i + \sum_{k=1}^{K} \alpha_k x_{kit} + u_{it} \tag{2}$$

Formula (2) gives the specification of the used fixed effect model. y_{it} presents the applied fertiliser of farmer i in round t. a_i stands for the corresponding fixed effect for farmer i. α_k is the estimated coefficient for variable k, whereas x_{kit} is the value of variable k from farmer i in round t. u_{it} is the error term of the estimation.

6. Results and discussion

Table 2 shows the results from estimating the fixed effect model. The control treatment is our reference point for the effects through the incentives. Except for the control treatment, all treatments show a shift in the second sequence at the 5% significant level. The shift for the control treatment is not significant. This indicates that each incentive has a significant negative influence on the amount of fertiliser used. Treatment A and treatment D with 47.9 kg and 47.7 kg respectively show the lowest shifts. Wald tests clearly indicates that treatment B and treatment C with 102.5 kg and 84 kg respectively have significantly higher shifts at a 5% level, whereas between these two pairs of treatments no significant difference is found. This shows that for the same effect on expected income, treatment B and C are significantly more effective. Therefore, it can be stated that the size of the shift depends on the incentives design. It seems for punishing, that a high probability of occurrence is more effective than a high magnitude, whereas for rewarding it is the other way round, holding effects on expected income constant. This means that our first hypothesis, i.e., *"H1: For the same effect on expected income, the effectiveness of an incentive is independent of its design."* has to be rejected. Based on the literature, this finding leads to several conclusions.

First, for punishment, we found the certain punishment, i.e., treatment B, to be more effective compared to the uncertain punishment with high magnitude, i.e., treatment D. A certain punishment leads to a stronger reaction than a higher and uncertain punishment, which implies that participants prefer to take the risk. Anderson and Stafford (2003) and Block and Vernon (1995) found that punishment magnitude has a higher influence on effectiveness than probability of occurrence. Kahneman (2011) states that low probabilities of occurrence lead to risk avoiding behaviour for losses. These findings contradict our research, since we found the opposite effect.

Second, for reward, treatment C with the high and uncertain magnitude is significantly more effective than the certain treatment A. Participants seems to prefer the risky treatment C over the certain treatment A. Volpp et al. (2008) found no difference in the effectiveness of magnitude or probability to occur for reward. Kahneman (2011) states that the low probability of occurrence leads to risk loving behaviour for gains. Especially the second finding is in line with our results.

Table 2: Estimation results of fixed effect model[a)]

	Coefficient	Standard Error	p-Value	
Shift, Control Treatment	26.6	(31.6)	0.399	
Shift, Treatment A	-47.9	(24.0)	0.040	*
Shift, Treatment B	-102.5	(23.8)	0.000	***
Shift, Treatment C	-84.0	(23.9)	0.000	***
Shift, Treatment D	-47.7	(23.7)	0.035	*
Learning, Sequence 1, All	9.95	(1.51)	0.000	***
Learning, Sequence 2, Control Treatment	4.18	(3.48)	0.230	
Sustainability, Treatment A	5.67	(3.47)	0.103	
Sustainability, Treatment B	14.33	(3.38)	0.000	***
Sustainability, Treatment C	2.02	(3.30)	0.541	
Sustainability, Treatment D	2.65	(3.22)	0.409	
Price, Sequence 1, All	0.031	(0.011)	0.006	**
Price, Sequence 2, Control Treatment	0.007	(0.014)	0.613	
Profit Difference, Treatment A	0.018	(0.002)	0.000	***
Profit Difference, Treatment B	0.011	(0.002)	0.000	***
Profit Difference, Treatment C	0.004	(0.002)	0.034	*
Profit Difference, Treatment D	-0.003	(0.002)	0.247	
Constant	200.6	(20.6)	0.000	***
Observations	3,800			
Adjusted R^2	0.63			

Source: Author's estimation

a) * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Third, comparing certain reward with certain punishment, i.e. comparing treatment A and treatment B respectively, we find punishing to be more effective. This is in line with the literature (Heath et al., 1999; Camerer et al., 1997; Fehr and Goette, 2007; Pope and Schweitzer, 2011). Also, in their meta-analysis regarding cooperation in social dilemmas, Balliet et al. (2011) found slightly higher but not significant differences in the effectiveness of punishment compared to reward. Additionally, Heath et al. (1999) state that people are loss averse, preconditioned there is a reference point. All of these statements are in line with our results that certain punishment is more effective than certain reward.

Fourth, if we hold expected effects on income constant, while increasing the magnitude of the incentives and reducing their probabilities to occur, reward (treatment C) becomes more effective than punishment (treatment D). It seems that the chance to get a high reward is a stronger motivator than the danger of having to receive a high punishment. This higher risk-lovingness for uncertain rewards compared to uncertain punishment can also be found in the literature (Kahneman, 2011).

For sustainability, a positive coefficient means that the average use of fertiliser rises in the second sequence. This means that the achieved drop in the use of fertiliser through the incentive gets lost again and the incentive is considered as being unsustainable. Treatment B has the only sustainability coefficient that is significantly higher than zero. Moreover, Wald tests indicate that at the 5% level it is significantly higher than treatment C and D, whereas for treatment A, this difference occurs at a significance level of 7.4 %. For treatments A, C and D the sustainability coefficients indicate no significant difference from zero at the 5% level; this indicates sustainability. Treatment B seems to be the only treatment where the shift through the incentive is not sustainable. Therefore, our second hypothesis, i.e., *"H2: For the same effect on expected income, the sustainability of an incentive is independent of its design."* has to be rejected.

Sustainability can be compared with iterated dilemmas. Balliet et al. (2011) found punishment to be insignificantly more sustainable in its effectiveness than reward. This contradicts our results, since certain punishment, i.e., treatment B, is the only treatment that is significantly less sustainable compared to the other treatments. If we combine the results of hypothesis one and two, we find that for creating a high and sustainable reduction for fertiliser use, treatment C, i.e., high but uncertain reward, is the preferable design for an incentive.

The profit difference is our measurement for efficiency in this experiment. For IDR one in profit difference, participants adopt their fertiliser use for 0.018 kg, 0.011 kg, 0.004 kg and 0.003 kg for treatment A, treatment B, treatment C and treatment D respectively. Since profit difference is our measure for efficiency, this also represents the order of efficiency for the different treatments in this experiment. Wald tests indicate different coefficients between all of these treatments. For the certain incentives, i.e., treatment A and treatment B, as well as for reward, i.e., treatment A and treatment C, participants show a significantly stronger reaction to the profit difference than to the uncertain or punishing incentives. Our third hypothesis, i.e., *"H3: For the same effect on expected income, the efficiency of an incentive is independent of its design."* has to be rejected.

Participants react strongest to the profit difference for certain and for rewarding incentives. This is in line with previous research that found that reward is more efficient than punishment (Rand et al., 2009; Sefton et al., 2007). In these papers, punishing can have a crucial effect on lowering the efficiency through generating costs. In our experiments, no such costs occur that underlines again our statement that reward is more efficient than punishment.

The learning effect in the first sequence is estimated simultaneously for all participants. The average amount of fertiliser used increases significantly by about 10 kg per round for the first sequence. The profit maximum would be reached with a fertiliser use of around 400 kg, depending on the price. Since most of the participants are clearly below that level of fertiliser use, it seems that maximising the profit influences the farmers in their decisions. For the second sequence, the learning effect is measured only for the control treatment. A not significant positive coefficient of 4.18 kg per round is estimated. This shows us that even though many participants are still clearly below the profit maximising use of fertiliser, the approach to the profit maximum slowed down or might have even stopped. It seems as if considerations other than maximising the profit influence farmers in their decision about the amount of fertiliser used.

The price effects are positive and they are significant for the first sequence. Participants seem to anticipate the price developments. For this experiment, a price raise of IDR 200 increases the fertiliser use by 6.3 kg for the first sequence and 1.4 kg for the control treatment in the second sequence.

7. Conclusion and outlook

The use of fertiliser in palm oil production creates negative externalities. If policymakers want to restrict such externalities, an effective, sustainable and efficient incentive system is desirable. The aim of this paper is to test differently designed incentives for policy measures on their effectiveness, sustainability and efficiency. For the first time, reward and punishment with different probabilities and magnitudes are simultaneously tested for policy measures. In doing so, we apply a business simulation game in a framed field experiment.

Results indicate differences between the incentives, depending on their design. Participants react strongest either to certain punishment or to uncertain reward, holding effects on expected income constant. For the later design, the reaction is also sustainable, making uncertain reward with high magnitude to be the most effective design. Furthermore, certain and rewarding incentives are significantly more efficient than uncertain and punishing incentives. To summarise, our findings suggest that adapting the magnitude and probability of occurrence of incentives can create a substantial difference for the effectiveness and efficiency of a policy measure. If policy focuses on ensuring a high and sustainable reduction in the use of fertiliser, our results suggest creating a high and uncertain reward. If policy focuses on being efficient, our results suggest creating a low but certain reward.

Using a business simulation game, we tried to achieve a high external validity. Nevertheless, there are implementations that may further improve external validity in the presented business simulation games. The production function, as well as prices for fruits or for fertiliser can be further specified to get a more realistic setting. In addition, for deeper insights regarding efficiency, the assumption of zero control costs can be reduced. It is likely that participants would react similarly if they are confronted with incentives for other measures, e.g., against erosion, reduction for pesticide use or also wildlife protection. However, these other measures are interesting extensions for further research, too.

The findings of this research are relevant for several groups. For public institutions it is interesting to know about effectiveness, sustainability and efficiency of incentive designs. This enables achieving public aims at relatively low costs. Also for farmers the results might be useful. Keeping our findings in mind, it may enable them a more objective view on policy measures. Furthermore, the present paper may be interesting for future research as we applied a business simulation game that enabled us to contribute to research gaps in the context of Indonesia.

Acknowledgements

This study is part of the Collaborate Research Centre 990. We thank the DFG (German Research Foundation) for funding and supporting this project. Moreover, we would like to thank our Indonesian colleagues, i.e., Dompak Napitupulu from the University of Jambi and Yusman Syaukat from Bogor Agricultural University, for their generous support and help. Furthermore, we thank Meike Wollni and Marcel Gatto from the University of Göttingen for their support in fieldwork and for sharing data with us. Last but not least, we want to thank our Indonesian enumerators, i.e. Zulfi Kamal and Fauzan Ishak. Without their patience especially regarding our understanding of Indonesian culture, data collection would not have been possible.

References

Anderson, L.R. and S.L. Stafford (2003): Punishment in a regulatory setting: experimental evidence from the VCM. In: Journal of Regulatory Economics 24 (1): 91–110.

Balliet, D.; Mulder, L. B. and P.A.M. Van Lange (2011): Reward, punishment, and cooperation: A meta-analysis. In: Psychological Bulletin 137 (4): 594–615.

Becker, G.S. (1974): Crime and punishment: An economic approach. In: Essays in the Economics of Crime and Punishment UMI: 1–54.

Block, M.K. and G.E. Vernon (1995): Some Experimental Evidence on Differences between Student and Prisoner Reactions to Monetary Penalties and Risk. In: The Journal of Legal Studies 24 (1): 123–138.

Camerer, C.; Babcock, L.; Loewenstein, G. and R. Thaler (1997): Labor Supply of New York City Cabdrivers: One Day at a Time. In: The Quarterly Journal of Economics (112): 407–441.

Fairhurst, T. and D. McLaughlin (2009): Sustainalbe Oil Palm Development on Degraded Land in Kalimantan. Washington, DC: World Wildlife Fund.

FAO (2005): Fertilizer use by crop in Indonesia. Rome: Food and Agriculture Organization of the United Nations (FAO); First version.

FAOSTAT (2013): http://faostat3.fao.org/faostat-gateway/go/to/download/Q/*/E. 18.October 2013.

Fehr, E. and L. Goette (2007): Do Workers Work More if Wages Are High? Evidence from a Randomized Field Experiment. In: The American Economic Review 97 (1): 298–317.

Fitzherbert, E.B.; Struebig, M.J.; Morel, A.; Danielsen, F.; Brühl, C.A.; Donald, P.F. and B. Phalan (2008): How will oil palm expansion affect biodiversity? In: Trends in ecology & evolution 23 (10): 538–545.

Harrison, G.W. and J.A. List (2004): Field experiments. In: Journal of Economic Literature 42 (4): 1009–1055.

Heath, C.; Larrick, R.P. and G. Wu (1999): Goals as Reference Points. In: Cognitive Psychology 38 (1): 79–109.

Herrmann, B.; Thöni, C. and S. Gächter (2008): Antisocial punishment across societies. In: Science 319: 1362–1367.

Hewitt, C.N.; MacKenzie, A.R.; Di Carlo, P.; Di Marco, C. F.; Dorsey, J. R.; Evans, M. et al. (2009): Nitrogen management is essential to prevent tropical oil palm plantations from causing ground-level ozone pollution. In: Proceedings of the National Academy of Sciences 106 (44): 18447–18451.

Ibanez, M. and P. Martinsson (2008): Can we do policy recommendations from a framed field experiment? The case of coca cultivation in Colombia. In: Working Papers in Economics 306. University of Gothenburg. School of Business, Economics and Law.

Jelsma, I.; Giller, K. and T. Fairhust (2009): Smallholder Oil Palm Production Systems in Indonesia: Lessons from the NESP Ophir Project. In: Plant Productin Systems, Plant Sciences Group, University Wageningen.

Kahneman, D. (2011): Thinking, Fast and Slow. New York: Farrar, Straus and Giroux.

Keller, M. and P.A. Matson (1994): Biosphere-Atmosphere Exchange of Trace Gases in the Tropics: Evaluating the Effects of Land Use Changes. In: Global Atmospheric-Biospheric Chemistry 48: 103-117.

Koh, L.P. and D.S. Wilcove (2008): Is oil palm agriculture really destroying tropical biodiversity? In: Conservation Letters 1 (2): 60–64.

Laumonier, Y.; Uryu, Y.; Stüwe, M.; Budiman, A.; Setiabudi, B. and O. Hadian (2010): Eco-floristic sectors and deforestation threats in Sumatra: identifying new conservation area network priorities for ecosystem-based land use planning. In: Biodiversity and conservation 19 (4): 1153–1174.

Levitt, S.D. and J.A. List (2007): What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World? In: The Journal of Economic Perspectives 21 (2): 153–174.

Palm, C. A.; Alegre, J.C.; Arevalo, L.; Mutuo, P.K.; Mosier, A.R. and R. Coe (2002): Nitrous oxide and methane fluxes in six different land use systems in the Peruvian Amazon. In: Global Biogeochemical Cycles 16 (4): 1–21.

Poitras, G. (1998): Spread options, exchange options, and arithmetic Brownian motion. In: Journal of Futures Markets 18 (5): 487–517.

Polinsky, A.M. and S.Shavell (1999): The economic theory of public enforcement of law. In: Working paper 6993. National Bureau of Economic Research. Cambridge.

Pope, D.G. and M. Schweitzer (2011): Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. In: The American Economic Review 101 (1): 129–157.

Rand, D.G.; Dreber, A.; Ellingsen, T.; Fudenberg, D. and M.A. Nowak (2009): Positive interactions promote public cooperation. In: Science 325: 1272–1275.

Reijnders, L. and M.A.J. Huijbregts (2008): Palm oil and the emission of carbon-based greenhouse gases. In: Journal of Cleaner Production 16 (4): 477–482.

Roe, B.E. and D.R. Just (2009): Internal and external validity in economics research: Tradeoffs between experiments, field experiments, natural experiments, and field data. In: American Journal of Agricultural Economics 91 (5): 1266–1271.

Sefton, M.; Shupp, R. and J.M. Walker (2007): The effect of rewards and sanctions in provision of public goods. In: Economic Inquiry 45 (4): 671–690.

Sekhon, G. S. (1995): Fertilizer-N use efficiency and nitrate pollution of groundwater in developing countries. In: Journal of Contaminant Hydrology 20 (3): 167–184.

Sutter, M.; Haigner, S. and M.G. Kocher (2010): Choosing the Carrot or the Stick? Endogenous Institutional Choice in Social Dilemma Situations. In: Review of Economic Studies 77 (4): 1540–1566.

Tiedemann, T. and U. Latacz-Lohmann (2013): Production Risk and Technical Efficiency in Organic and Conventional Agriculture - The Case of Arable Farms in Germany. In: Journal of Agricultural Economics 64 (1): 73–96.

Tute, C. (2005): An outcome-based payment scheme for the promotion of biodiversity in the cultural landscape. In: Integrating Efficient Grassland Farming and Biodiversity: 36-39.

UNCTAD (2013): Palm oil: www.unctad.info/en/Infocomm/AACP-Products/Palm-oil/. 21.October 2013.

Veldkamp, E. and M. Keller (1997): Nitrogen oxide emissions from a banana plantation in the humid tropics. In: Journal of Geophysical Research: Atmospheres (1984–2012) 102: 15889–15898.

Veldkamp, E.; Purbopuspito, J.; Corre, M.D.; Brumme, R. and D. Murdiyarso (2008): Land use change effects on trace gas fluxes in the forest margins of Central Sulawesi, Indonesia. In: Journal of Geophysical Research: Biogeosciences 113(G2).

Volpp, K.G.; John, L.K.; Troxel, A.B.; Norton, L.; Fassbender, J. and G. Loewenstein (2008): Financial incentive–based approaches for weight loss. In: JAMA 300 (22): 2631–2637.

Wilcove, D.S. and L.P. Koh (2010): Addressing the threats to biodiversity from oil-palm agriculture. In: Biodiversity and conservation 19 (4): 999–1007.

Appendix A: Instructions

The experiment took place in a public building, e.g., a school, a sports hall, an administration building, or in the house of the village head. Depending on the size of the village, we invited between 10 and 18 randomly chosen farmers. The experiment started either after lunch or after the evening prayer. When the farmers arrived, they were identified and interviewed. Here, personal as well as economic information from the farmers was gathered. For the experiment, the farmers sat on the floor or, if available, on chairs with tables. To avoid chatting among the farmers, we placed them in distance to each other. One researcher and three or four enumerators assured a smooth conduction of the experiment. For the experiment, we used questionnaires, one for each farmer, including a fertiliser-yield table (Figure A1), an example sheet (Figure A3) and one decision sheet for each of the ten round (Figure A2). Furthermore, to support our explanation, we used posters that were equal to the questionnaire sheets.

The Experiment starts:

Welcome to our experiment and thank you for coming. *We introduced ourselves, our research project and our home university*. This experiment simulates a palm oil plantation being owned and operated by you. The results from this experiment are of academic interest. If you complete it you will earn a shopping voucher for a local shop. The value of the shopping voucher depends on the decisions you have to make during this experiment, so listen carefully. Whenever you have a question, do not hesitate to ask. The experiment will take approximately one hour and will start immediately.

We distributed the questionnaires to the farmers and hung up posters to support the explanations (see Figure A1, Figure A2 and Figure A3).

We start explaining the simulated situation to the farmers: Imagine you have a palm oil plantation of about one hectare. Your plantation is already established and in production. You have to decide how much fertiliser you want to use, which will affect your yield. Then, you will sell this yield, pay for the fertiliser and, additionally, you might receive some public money. So, you will end up with a profit. You will receive a shopping voucher in the value of 5% of this profit. Now, we will explain you the experiment in detail.

This experiment simulates a palm oil production for 10 years, where each year equals one round. At the beginning of each year, you have to decide how much fertiliser you want to use. This is the only decision you have to make during the year. You are allowed to use between 0 and 1,000 kg fertiliser, but only in multiples of 10. The fertiliser used is the only determinant for the yield you receive. This means that the experiment does not consider diseases, pests, weather conditions etc. We are aware that this is not very realistic, but for the purpose of this experiment, this assumption is necessary. To make the fertiliser decision easier for you, we put a fertiliser-yield table (Figure A1) in your questionnaire showing the yield for a certain amount of fertiliser.

After you decided how much fertiliser you will use, the price for your yield is ascertained. For the first year, this price is IDR 1,800 per ton palm oil bunches. To generate the price for each of the following years, you have to draw a ball from a bag with three green and tree red balls. If you draw a green ball, the price rises by IDR 200 compared to the previous year, whereas the price falls by IDR 200 if you draw a red ball.

When the price is known, an enumerator will calculate your profit for the actual year. Thus, the enumerator declares your yield and, based on the price, calculates your revenue. Then the costs for fertiliser, which are at IDR 10 per kg, are subtracted. Additionally, there is an unconditional public subsidy of IDR 10,000 per year, which is added. Later on, this public intervention can change. If the profits for all present farmers in a year are calculated, we start with the next year. Again, you decide how much fertiliser you want to use. At the end of the experiment, all profits are added up and you will receive 5% of these profits in the form of a shopping voucher for a local shop.

Take a look at the example provided in your questionnaire (Figure A3). There is a fertiliser use of 100 kg. With a look in the fertiliser-yield table, which is implied in your questionnaire, you can see that this results in a yield of 14.5 tons. Furthermore, you can see that the price in the previous round is IDR 1,200 per ton. Since a green ball is drawn in this example, the price increases for IDR 200 to IDR 1,400 per ton. A yield of 14.5 tons multiplied by a price of IDR 1,400 per ton results in a revenue of IDR 20,300. To account for the costs of fertiliser, 100 kg multiplied by a price of IDR 10 per kg are deducted. Additionally, there is a subsidy of IDR 10,000. In total, this results in a profit of IDR 29,300 for this year. 5% of this amount, i.e. IDR 1,465, would have been added to your shopping voucher.

To check if you understood the experiment, we have some control questions. *We asked the following questions one by one to the audience; if necessary, we repeated our explanation until everyone had understood the answer.*

- What kind of plantation do you cultivate in the experiment?
- What is the size of your plantation?
- How many years of palm oil production simulates this experiment?
- If you use 50 kg fertiliser, how much yield do you get?
- If you use 150 kg fertiliser, how much yield do you get?
- If you want to achieve a yield of 20.1 tons, how much fertiliser do you need?
- If the price in the previous year was IDR 1,600 per ton, and you draw a green ball, what is the new price?
- If the price in the previous year was IDR 2,000 per ton, and you draw a red ball, what is the new price?
- What are the costs for 170 kg of fertiliser?
- What amount of subsidy do you receive each year?
- Can the subsidy change during the experiment?

Please ask further questions! If there are no questions left, we can start with the experiment. Please make your own decisions and do not talk to someone except the enumerators during the experiment. If you talk to other farmers during the experiment, the data is not useful for us and therefore you will be excluded from the experiment.

We formed small groups of three to five farmers with one enumerator during the experiment. This eased conduction of the experiment and, if necessary, the communication between the farmers and the enumerator. We conducted round 1-5. After the fifth round, the procedure differed between the treatments as follows:

For the control treatment:
As there is no difference for the control treatment between round 1-5 and round 6-10, the experiment continues without any interruption.

For the treatment A:
The subsidy of IDR 10,000 is cancelled. Instead, there is a new public intervention program, which aims to reduce the use of fertiliser. It is a compensation payment of IDR10,000 for everyone who uses 120 kg or less fertiliser per ha.

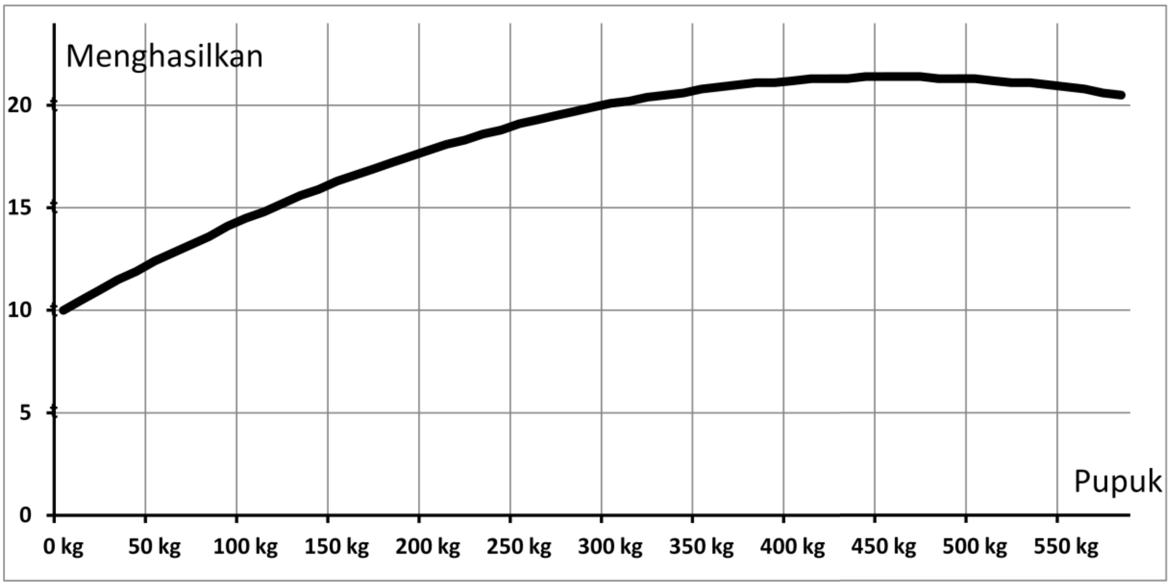
For the treatment B:
In addition to the subsidy of IDR 10,000 per year, there is a new law. It aims to restrict the use of fertiliser. It says that there is a punishment of IDR 10,000 for everyone who uses 130 kg fertiliser or more.

For the treatment C:
The subsidy of IDR 10,000 is cancelled. Instead, there is a new public intervention program, which aims to avoid negative externalities generated by the use of fertiliser. There will be a compensation payment of IDR 100,000 for everyone who avoids such negative externalities. If you use 120 kg or less fertiliser, you have a 10% chance to avoid such externalities. To determine if the incentive takes place or not, we use a bag with nine blue and one yellow ball. If you draw a blue ball, you were not able to avoid the externalities and you do not receive the compensation payment. If you draw the yellow ball, you where able to avoid these negative externalities and you receive the compensation payment of IDR 100,000.

For the treatment D:
In addition to the subsidy of IDR 10,000 per year, there is a new law. It aims to restrict the use of fertiliser. It says that there is a punishment of IDR 100,000 for everyone who uses 130 kg fertiliser or more. The control for this law is not perfect. There is a chance of only 10% to get caught. To determine if you get caught, we use a bag with nine blue and one yellow ball. If you draw a blue ball, you were not caught and you do not get any punishment. If you draw the yellow ball, you were caught and you have to pay a punishment of IDR 100,000.

We conducted round 6 – 10. Subsequently, we summed up the profits and gave the corresponding shopping vouchers to the participants.

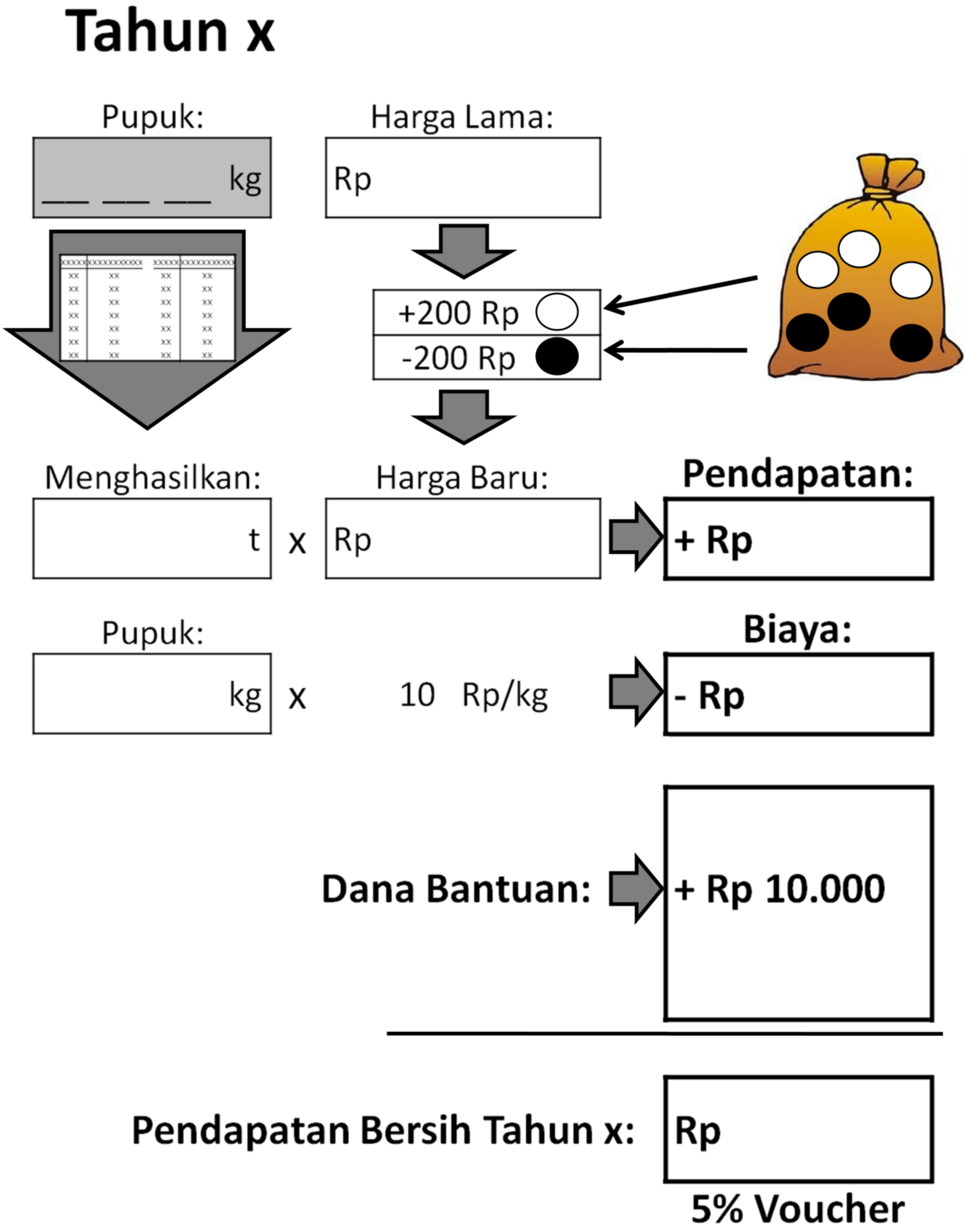
Figure A1: Fertiliser-yield table[a)]



Pupuk	Mengha-silkan	Pupuk	Mengha-silkan	Pupuk	Mengha-silkan	Pupuk	Mengha-silkan
0 kg	10,0 t	150 kg	16,3 t	300 kg	20,1 t	450 kg	21,4 t
10 kg	10,5 t	160 kg	16,6 t	310 kg	20,2 t	460 kg	21,4 t
20 kg	11,0 t	170 kg	16,9 t	320 kg	20,4 t	470 kg	21,4 t
30 kg	11,5 t	180 kg	17,2 t	330 kg	20,5 t	480 kg	21,3 t
40 kg	11,9 t	190 kg	17,5 t	340 kg	20,6 t	490 kg	21,3 t
50 kg	12,4 t	200 kg	17,8 t	350 kg	20,8 t	500 kg	21,3 t
60 kg	12,8 t	210 kg	18,1 t	360 kg	20,9 t	510 kg	21,2 t
70 kg	13,2 t	220 kg	18,3 t	370 kg	21,0 t	520 kg	21,1 t
80 kg	13,6 t	230 kg	18,6 t	380 kg	21,1 t	530 kg	21,1 t
90 kg	14,1 t	240 kg	18,8 t	390 kg	21,1 t	540 kg	21,0 t
100 kg	14,5 t	250 kg	19,1 t	400 kg	21,2 t	550 kg	20,9 t
110 kg	14,8 t	260 kg	19,3 t	410 kg	21,3 t	560 kg	20,8 t
120 kg	15,2 t	270 kg	19,5 t	420 kg	21,3 t	570 kg	20,6 t
130 kg	15,6 t	280 kg	19,7 t	430 kg	21,3 t	580 kg	20,5 t
140 kg	15,9 t	290 kg	19,9 t	440 kg	21,4 t	590 kg	20,4 t

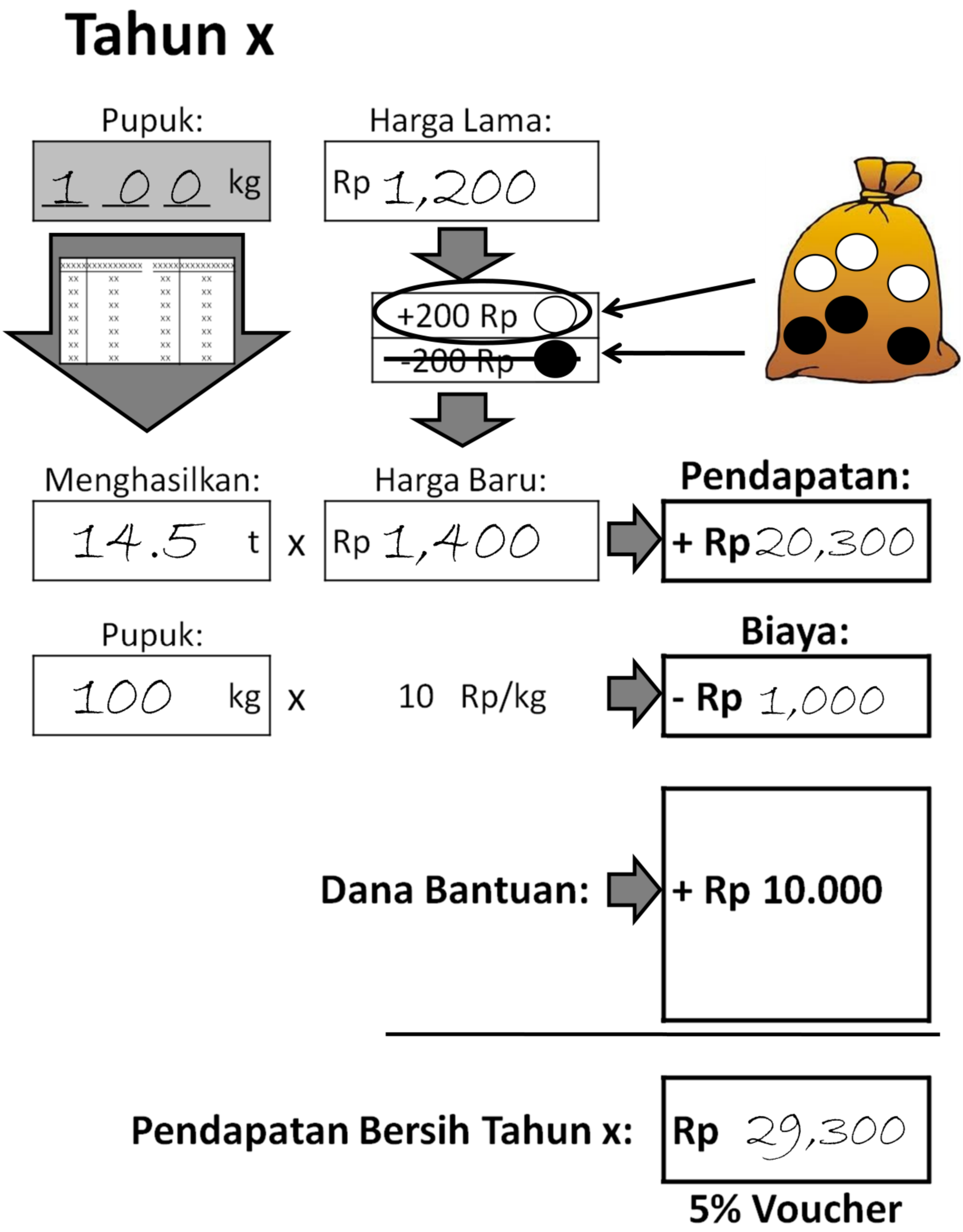
a) Translation: Menghasilkan = production of palm oil bunches; Pupuk = fertiliser

Figure A2: Decision sheet[a)]



a) Translation: Tahun x = year x; Pupuk = fertiliser; Harga Lama = old price; Menghasilkan = production of palm oil bunches; Harga Baru = new price; Pendapatan = revenue; Biaya = cost; Dana Bantuan = Relief fund; Pendapatan Bersih Tahun x = net income year x;

Figure A3: Example of decision sheet[a)]



a) Translation: Tahun x = year x; Pupuk = fertiliser; Harga Lama = old price; Menghasilkan = production of palm oil bunches; Harga Baru = new price; Pendapatan = revenue; Biaya = cost; Dana Bantuan = Relief fund; Pendapatan Bersih Tahun x = net income year x;