GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

GlobalFood
www.uni-goettingen.de/globalfood

RTG 1666 GlobalFood

Transformation of Global Agri-Food Systems:
Trends, Driving Forces, and Implications for Developing Countries

University of Goettingen

GlobalFood Discussion Papers

No. 140

Contract compliance under biased expectations:
Evidence from an experiment in Ghana

Sabine Fischer
Kerstin Grosch

March 2020

RTG 1666 GlobalFood · Heinrich Düker Weg 12 · 37073 Göttingen · Germany
www.uni-goettingen.de/globalfood

ISSN (2192-3248)

Suggested Citation:

Fischer, S., K. Grosch (2020). Contract compliance under biased expectations: Evidence from an experiment in Ghana. GlobalFood Discussion Paper 140, University of Goettingen. http://www.uni-goettingen.de/de/213486.html.

Contract compliance under biased expectations: Evidence from an experiment in Ghana*

Sabine Fischer †1 and Kerstin Grosch ‡2

[1]University of Goettingen
[2]Institute for Advanced Studies (IHS)

March 1, 2020

Abstract

Contract compliance is key for economic growth. However, determinants affecting contract breach are not yet well understood. In this paper, we focus on contract situations with a potential hold-up problem, such as contract farming agreements which are prevalent in many developing countries. We examine if agents' payoff expectations serve as a reference point affecting (non-)compliant behavior by inducing a subjective loss when the agent compares the realized payoff and the expected payoff from the contract. Results from our lab experiment in Ghana indicate that overconfident agents, i.e., agents with relatively high payoff expectations, breach more often than underconfident agents, i.e., agents with relatively low payoff expectations. Moreover, more pronounced individual loss aversion amplifies the effect of subjective losses on contract breach. In a treatment, we manipulate agent's overestimation exogenously and use it as an instrument to demonstrate that the reported effects are causal.

JEL codes: C91, D03, D81, D84, J41, O12
Keywords: contract compliance, overconfidence, reference-dependent preferences

---

*For helpful comments we thank Marcel Fafchamps, Claus Ghesla, Simone Häckl, Johannes Haushofer, Marcela Ibanez, Abdul Qadir Ibrahim, Emmanuel Peterle, Holger A. Rau, Gerhard Riener, Michala Iben Riis-Vestergaard, Florian Spitzer, Stefan Trautmann, Angelino Viceisza, Meike Wollni, Christina Felfe's and Conny Wunsch's chair, and our colleagues of the research group GlobalFood. Furthermore, we acknowledge funding from the German Research Foundation (DFG), grant number RTG 1666 (GlobalFood). Thanks goes also to the Agriculture Department at the Kwame Nkrumah University of Science and Technology (KNUST) and our research assistants in Accra und Kumasi for support in the field.

†Heinrich-Düker-Weg 12, 37073 Goettingen (Germany)

‡Corresponding author, Josefstädter Str. 39, 1080 Vienna (Austria), *E-mail: grosch@ihs.ac.at*

1 Introduction

Encouraging contract compliance can be challenging, especially when legal enforcement mechanisms do not function properly (Chemin, 2012). Particularly in developing countries, legal systems are often absent, corrupt or too slow to be usable (Dixit, 2003). Hence, contract conclusion is a risky endeavor in terms of post-contractual opportunistic behavior despite potentially high returns on investment (Klein et al., 1978). Positively expressed, contracts can be an important tool to stimulate growth (Reardon et al., 2009). Therefore, learning more about determinants of compliant behavior in developing countries is essential. In this paper, we investigate the effect of biased payoff expectations on agents' compliant behavior in contract situations.

It has been shown that people evaluate payoffs as gains and losses rather than as a detached state of wealth (Kahneman and Tversky, 1979). There is a class of theories developed upon the assumption that agents use their payoff expectations as reference payoffs; agents are loss averse around their expected payoff, and losses are more painful than equal-sized gains are pleasurable (e.g., Bell, 1985; Loomes and Sugden, 1986; Gul, 1991; Koeszegi and Rabin, 2006). Utility is then dependent not only on a realized payoff but also on the payoff expectations, that is the reference payoff, to which the realized payoff is compared. When the realized payoff falls short of the expected payoff, the agent may sense a subjective loss which may alter compliant behavior ex post.

Our study is based on the idea that payoff expectations may be biased. There is extensive evidence that people often do a poor job of processing probabilities, fail in assessing their performance, and may form overconfident performance expectations (e.g., Tversky and Kahneman, 1975; Kruger and Dunning, 1999; Moore and Healy, 2008; Benoît et al., 2015). In contract situations, payoff expectations may stem from the biased assessment of one's own performance. For example, in agriculture a farmer's payoff depends crucially on the skills and effort exerted. When farmers overestimate their own performance resulting in a shortfall in expected production, it may directly translate into unmet payoff expectations.

Our main hypothesis is that contract breach increases agent's overestimation in contract situations in which the payoff depends on performance. The mechanism that we test is expectation-based, reference-dependent preferences. Furthermore, we examine two moderators that may bolster reference-dependent behavior. One moderator may be the level of individual loss aversion. Along the lines of the model on reference-dependent preferences by Koeszegi and Rabin (2006), subjective losses may resonate more intensely the higher the level of individual loss aversion. To counteract the relatively high subjective loss, overconfident agents with a pronounced loss aversion might try to increase overall utility by selling at a more lucrative market. Another moderator that we test is the nature of the environment which may be deterministic, i.e., production shocks cannot occur, or non-deterministic, i.e., production shocks can occur. Production shocks are prevalent in many businesses. In the farming context, weather shocks (positive or negative) may affect production outcomes. In such an environment, agents can attribute their performance failure to external factors according to their ex-ante beliefs and utilize the shock to legitimize contract breach while keeping up their self-image (e.g., Blanton et al., 2001; Köszegi, 2006; Dana et al., 2007; Grossman and Owens, 2012). Therefore, breaching might be facilitated in a non-deterministic environment compared to a deterministic environment. We explain these mechanisms in further detail using a theoretical framework in the following section.

We test the hypotheses in a lab experiment with students in Ghana. Since our research questions are particularly interesting for developing economies, we consider our subject pool from Ghana as an asset of this study.[1]

The experimental design is based on a contract situation in the agricultural sector. In many developing countries, agriculture is the largest sector in terms of labor force engaged. In Ghana, agriculture contributes around 55% to the GDP and a similar share of about 55% of the labor force is engaged in this sector (FAO, 2019). So-called "contract farming" is prevalent in developing countries in which farmers lack resources.[2] Contract farming is usually a hold-up situation and appears in different forms. Our study design is based on a rather typical contract-farming situation where a buyer (principal) provides money to a seller (agent) who, in return, promises to sell a certain number of goods to a buyer in the future for a predetermined price.

To translate the contract farming situation into the lab, we use a multi-stage investment game where an agent and a principal conclude a contract with an inherent hold-up problem. In the final stage of the experiment, agents decide to comply with the contract and sell the amount they have offered to the principal. This is a continuous decision and agents freely decide how much they ultimately want to sell to the principal. To identify whether the effect from overestimation is causal, we manipulate expectations through an experimental treat-

[1]The majority of laboratory experiments in economics and psychology are conducted with participants from WEIRD (Western, educated, industrialized, rich, and democratic) societies. Experimental findings indicate considerable behavioral differences among societies in diverse domains such as analytic reasoning, fairness, and cooperation (Henrich et al., 2010; Jones, 2010). We are aware that students may not be representative of Ghana's population and for smallholder farmers in particular, who can be assumed to be poorer, less educated, and more vulnerable to income shocks than students (Haushofer and Fehr, 2014). However, our research questions required a rather complicated experimental design that would have been difficult to conduct with farmers. We expect that we may rather underestimate effects since research suggests that poor people may be particularly loss averse compared to wealthier people (Tanaka et al., 2010).

[2]For further reading on contract farming and its welfare effects, take a look at Wang et al. (2014), Ton et al. (2018), and Bellemare and Bloem (2018).

ment in which we use an anchor to inflate performance guesses (Tversky and Kahneman, 1975).[3] We elaborate on the experimental design in section 2.2.

We establish the following results. Overconfident agents are more likely to breach a contract compared to non-overconfident agents. This result is driven by underconfident agents who breach less or even "over-fulfill" contracts, i.e., they sell more than they had initially promised in the contract. We can confirm the hypothesis that loss aversion increases contract breach for overconfident agents. We find that results on contract breach remain significant once overestimation is instrumented by the anchor treatment, implying that the effect is causal. Results suggest that a non-deterministic environment makes breaching more likely. Apart from our main experimental results, we find further convincing evidence of reference-dependent behavior. Data analysis from an ex-post questionnaire reveals that agents with similar final payoffs are less satisfied when they were overconfident compared to non-overconfident agents. Moreover, overconfident agents state more often that they felt disappointed when informed about their realized payoff compared to non-overconfident agents.

Previous studies on reference points have predominantly focused on ex-ante behavioral change. For instance, it has been shown that customers adjust their consumption behavior (Huang and Liu, 2019), people pay their taxes (Alm et al., 1992), workers increase their effort level (Abeler et al., 2011), traders reduce their valuation on endowed goods (Marzilli Ericson and Fuster, 2011), golf players invest more focus when going for par (Pope and Schweitzer, 2011), and gamblers gear their lottery choices (Kahneman and Tversky, 1979) to mitigate anticipated losses. There are few studies that focus on behavioral change *after* experiencing a subjective loss, as in our study. Reference points can be exogenous, as in Abeler et al. (2011). Several studies demonstrate that people adjust their reference points dynamically according to new circumstances, such as Gill and Prowse (2012) in a competitive situation and Post et al. (2008) in the game show "Deal or No Deal." We hypothesize that reference points can also form as the result of individual overconfidence which is also a novel feature of our study.

In the contract literature, there are few studies that focus on the consequences of reference points for contract compliance. Hart and Moore (2008) demonstrate theoretically and Fehr et al. (2011) show experimentally that agents breach less under a rigid contract, i.e., a contract in which the price for the agent's output is fixed ex ante, than under a flexible contract, i.e., the price is not fixed ex ante but can vary within a certain range. Under flexible contracts, agents engage in costly punishment if the principal pays a lower price

[3] This approach is comparable to Hill et al. (2012) who examine whether observing peers alters the decision to reciprocate in a trust game. By using a treatment, they instrument their explanatory variable of interest and justify therewith that their reported effect is indeed causal.

than expected. The sanctioning party forgoing economic payoff to enforce contracts can be called a “second party.” Second-party enforcement can be driven by existing social norms to comply, repeated contract interactions, and a potential loss in reputation (Klein, 2000; Fehr and Fischbacher, 2004). In contrast, first-party enforcement is based on a party’s internal value system and individual preferences such as guilt aversion, fairness concerns, and social preferences affecting utility of breaching and ultimately the decision to comply to a contract. There is extensive empirical evidence from behavioral economics that first-party as well as second-party enforcement mechanisms play a crucial role for contract compliance (e.g., Fehr et al., 2002; Charness and Dufwenberg, 2006; Koszegi, 2014). We incorporate these insights from behavioral economics into our study and control for individual preferences such as guilt aversion and risk preferences.

Closest to our paper is a recent study by Garbarino et al. (2019) who examine how expected monetary outcomes alter an individual’s likelihood of lying to make profit. The authors manipulate payoff expectations exogenously by varying the probability of high and low payoff outcomes. They show that subjects lie more ex post when payoff expectations had been higher ex ante and explain the results with reference-dependent behavior and loss aversion. While the study of Garbarino et al. (2019) and our study both examine how initial payoff expectations affect opportunistic behavior, our study differs in at least three important aspects. First, we attempt to experimentally investigate the effect of an *endogenous* reference payoff from overconfidence on ex-post behavioral changes and not from exogenously-varied payoff expectations. Second, rather than examining whether the observed behavior is consistent with a theory of loss aversion on an aggregate level, our design allows us to measure and test the effect of loss aversion on an individual level. Third, we look at a contract situation with a preceding promise.

We contribute to the existing literature on reference-dependent preferences, contract behavior, and overconfidence in different ways. First, we formally establish and test how overestimation affects compliant behavior in contract situations. Second, we contribute to the literature on the formation of reference points and provide evidence that they can emerge endogenously as a result of one’s own performance (mis-)judgment. Third, we account for individual heterogeneity and experimentally demonstrate that individual loss aversion indeed moderates the effect of a reference payoff. Fourth, we provide evidence of reference-dependent preferences in a non-WEIRD society.

The remainder of the paper proceeds as follows. We deduce propositions from a theoretical framework and explain the overall study design in section 2, followed by the presentation of the results in section 3, and a discussion in section 4. Section 5 concludes with a summary of the results and potential policy implications.

2 Study design

2.1 Theoretical Framework

In this section, we suggest a theoretical framework to deduce propositions for our contract situation with a hold-up problem. Similar to Garbarino et al. (2019), we use a model in which utility is reference-dependent and losses loom larger than gains (e.g., Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). There are two parties, an agent (seller) and a principal (buyer) who can agree on a mutual contract. Once the contract is formed, the agent can decide whether to breach or to comply with the contract. Our model only considers the agent's decision after the contract is concluded. We assume that an agent's utility consists of the following additive components. Firstly, the agent gains standard consumption utility for the realized payoff π_R which results from the payoff π_0 according to the contract and an additional monetary surplus π_{br} from breaching. Secondly, the utility is affected by reference-dependent utility gains or losses, indicated by the functions g() and h() respectively, which depend on the difference between the realized payoff π_R and the expected payoff π_e. When the expected payoff exceeds the realized payoff, the agent is overconfident and underconfident vice versa. The third component that affects the agent's utility is the intrinsic moral costs of breaching $\Omega()$ as a function of the agent's social preferences ω, the extent of breaching π_{br}, and the environment n.

The agent's utility function can be written as:

$$\begin{aligned} U(\pi_R|\pi_{br}) = u(\pi_0 + \pi_{br}) + g(\pi_0 + \pi_{br} - \pi_e) I_g(\pi_e < \pi_0 + \pi_{br}) \\ -h(\lambda, (\pi_e - \pi_0 + \pi_{br})) I_h(\pi_e > \pi_0 + \pi_{br}) - \Omega(\pi_{br}, \omega, n), \end{aligned} \tag{1}$$

where

π_R	=	agent's realized payoff,
π_0	=	agent's payoff according to the agreed upon contract,
π_{br}	=	agent's additional payoff from breaching the contract,
π_e	=	agent's expected payoff,
λ	=	agent's individual loss aversion ($\lambda \geq 1$),

$\Omega(\pi_{br}, \omega, n)$	=	moral costs from breaching the contract which are increasing in agent's social preferences ω, and the amount breached π_{br}. The variable n is binary and indicates whether the environment is deterministic or non-deterministic. We assume moral costs to be lower when the environment n is non-deterministic compared to a deterministic environment,
$g()$	=	utility gains, where $g'(\pi_R - \pi_e) > 0$, and $g''(\pi_R - \pi_e) < 0$,
I_g	=	indicator function that takes the value 1 if the argument in brackets is true (agent experiences a gain) and the value 0 if the argument is false (agent does not experience a gain),
$h()$	=	utility loss depending on the agent's individual loss aversion parameter λ, where $h'(\pi_e - \pi_R) > 0$, and $h''(\pi_e - \pi_R) < 0$,
I_h	=	indicator function that takes the value 1 if the argument in parentheses is true (agent suffers a loss) and the value 0 if the argument is false (agent does not suffer a loss).

Following the concept of loss aversion, we assume that $g()$ and $h()$ are not symmetric and that losses weigh larger than gains whereby the steepness of the loss function $h()$ increases with the individual loss aversion parameter λ.

In equation (1), we formalized the realized payoff π_R which is the payoff *after* the agent's decision to breach or to comply to the contract. *Before* this decision, an overconfident agent experiences a subjective loss $h()$ unambiguously since the expected payoff π_e is lower than the payoff from the contract π_0. Similarly, an underconfident agent experiences a subjective gain $g()$. Agents who estimate their payoff correctly experience neither a subjective loss nor a subjective gain. By breaching the contract, agents not only increase the consumption utility by increasing their realized payoff but they also alter the level of a subjective loss $h()$ (or a subjective gain $g()$) and, therefore, they can decrease their subjective loss (or increase their subjective gain) the higher the level of the breach. An overconfident agent may even ascend from the initial loss domain to the gain domain when the agent breaches to a high enough extent. However, breaching to the highest extent may not be utility maximizing since agents also face moral costs $\Omega()$ from breaching the contract.

In the following, we deduce propositions from the theoretical framework regarding varying payoff expectations, individual loss aversion and the type of the environment: deterministic or non-deterministic.

We assume that the agent breaches the contract when the utility from breaching is higher than the utility from compliance:

$$u(\pi_0+\pi_{br})+g(\pi_0+\pi_{br}-\pi_e)I_g(\pi_e<\pi_0+\pi_{br})-h(\lambda,(\pi_e-\pi_0+\pi_{br}))I_h(\pi_e>\pi_0+\pi_{br})-\Omega(\pi_{br},\omega,n)$$
$$>u(\pi_0)+g(\pi_0-\pi_e)I_g(\pi_e<\pi_0)-h(\lambda,(\pi_e-\pi_0))I_h(\pi_e>\pi_0) \tag{2}$$

First, we examine the effect of a marginal increase of over- and underestimation. For this, we hold the breaching level constant at $\bar{\pi}_{br}$. An overconfident agent breaches the contract when the following condition is met:[4]

$$u(\pi_0+\bar{\pi}_{br})-h(\pi_e-\pi_0-\bar{\pi}_{br})-\Omega(\bar{\pi}_{br},\omega,n)-u(\pi_0)+h(\pi_e-\pi_0)>0 \tag{3}$$

Let us assume payoff expectations are marginally higher by ϵ ($\epsilon>0$), i.e., one agent is more overconfident than another. The resulting condition is given by:

$$u(\pi_0+\bar{\pi}_{br})-h(\pi_e+\epsilon-\pi_0-\bar{\pi}_{br})-\Omega(\bar{\pi}_{br},\omega,n)-u(\pi_0)+h(\pi_e+\epsilon-\pi_0)>0 \tag{4}$$

To demonstrate the effect of more pronounced overestimation, we compare the utility loss functions of the inequalities (3) and (4):

$$h(\pi_e+\epsilon-\pi_0-\bar{\pi}_{br})-h(\pi_e+\epsilon-\pi_0)>h(\pi_e-\pi_0-\bar{\pi}_{br})-h(\pi_e-\pi_0) \tag{5}$$

Since the gain-loss utility function is concave the left side of inequality (5) is always larger than its right side. This means that a marginal increase in overestimation is reducing the utility gain from breaching a fixed amount $\bar{\pi}_{br}$. Or in other words, the more an agent's estimated payoff falls short from the realized payoff, the smaller the marginal utility gain is from breaching.

Similarly, underconfident agents face lower utility gains from breaching with increasing underestimation since the gain function is also concave. Therefore, we expect the probability of breaching the contract to decrease with increasing underestimation.

Proposition I: *Overconfident agents are less likely to breach the contract with increasing overestimation. Underconfident agents are less likely to breach the contract with increasing underestimation.*

[4] We simplify this equation and the following equation (4) and exclude the case of overconfident agents who breach a lot and end up in the gain domain for reasons of brevity and easier notation/understanding. However, the line of argument and inferences we make are not affected by this simplification.

Above, we compared utilities between complying to and breaching the contract in a static fashion. Next, we will discuss expectations for the level of breaching. For this purpose, we take the first derivative with respect to π_{br}. The optimal level of breaching π^*_{br} is reached when the marginal benefits equal the marginal costs of breaching. Since the utility function is discontinuous, we regard the derivation for the case of overestimation (loss function) first, and infer from that for the case of underestimation (gain function) for the sake of brevity:

$$u'(\pi_0 + \pi^*_{br}) - h'(\lambda, (\pi_e - \pi_0 + \pi^*_{br})) = \Omega'(\pi^*_{br}, \omega, n) \tag{6}$$

A higher level of overestimation, i.e., rising expected benefits π_e, decreases the argument on the left side c.p. because it increases the subjective loss and, therefore, increases h'() due to its functional form. The optimal breaching level π^*_{br} has to be higher for agents with relatively high overestimation compared to agents with relatively low overestimation to fulfill the maximization condition since $h'()$ is decreasing with an increasing π_{br}.
We expect that the individual loss aversion λ serves as a weight for the subjective loss, i.e., the more pronounced λ, the steeper the utility curve in the loss domain and the more pronounced the subjective loss $(\pi_e - \pi_0 + \pi^*_{br})$. Hence, we expect that the effect of overestimation on breaching is conditional on individual loss aversion.
Similar to the reasoning about overestimation, the optimal breaching level of underconfident agents must be lower, the higher their underestimation.

Proposition II: *a) Agents who breach the contract, breach contracts to a higher extent with increasing overestimation, conditional on their loss aversion. b) Agents who breach the contract, breach to a lower extent with increasing underestimation.*

The moral costs of breaching $\Omega()$ may depend on the type of environment. A non-deterministic environment provides wiggle room and the agent can shift own accountability for the breaching action to the environmental conditions, agents can avoid negative self-signals from failing to meet expectations, and they do not have to update their self-concept (Sackeim and Gur, 1979; Kőszegi, 2006; Dana et al., 2007; Mazar et al., 2008). Hence, the moral costs $\Omega(\pi_{br}, \omega, n)$ may be reduced within a non-deterministic environment with an exogenous production shock compared to a non-deterministic environment. Therefore, we expect that agents in our shock condition (non-deterministic environment) will be more likely to breach the contract and will breach to a higher extent than sellers in the no-shock condition (deterministic environment).

Proposition III: *Agents are more likely to breach in the shock condition (non-deterministic environment) and will breach the contract to a higher extent than agents in the no-shock condition (deterministic environment).*

2.2 Experimental design and survey

To test the above-stated propositions, we conducted an experiment consisting of a short pre-survey, elicitation of individual preferences, and an investment game which is followed by an ex-post questionnaire.[5]

The experiment was programmed on tablets with the software "oTree" (Chen et al., 2016). Sessions lasted approximately 120 minutes and students were paid, in total, about GHS 44 ($11), including a show-up fee of GHS 20. Payments varied between GHS 22 and GHS 90.

Investment game

At the beginning, subjects are matched in dyads and are randomly assigned to the role of either a seller or a buyer. Sellers produce goods and buyers can invest in a seller's production to enhance the goods' value. Sellers exert a real-effort task, simulating the production process, and state a performance guess afterwards. Based on their guess, sellers make an offer for a certain number of goods to sell to the buyer. A contract is concluded when the buyer accepts the offer. Then, sellers are informed about their expected payoff as well as their realized payoff if they complied with the contract. Sellers can then decide how many goods to sell to the buyer eventually, which can result in compliance, breaching or even "over-fulfilling" by selling as much output as was promised, less output or even more output than was promised. By breaching, agents sell more of their output to the lucrative alternative market, and overconfident sellers can offset the divergence between the realized and expected payoff.

In the following, we describe the different steps in the contract phase in more detail. Steps 2 to 5 are illustrated in Figure 1. All participants learn about all the different steps of the experiment before making their decisions.

1. **Real-effort task and performance guess:** Sellers are confronted with 25 puzzles with four possible answers each. We explain to subjects that correctly solved tasks correspond to produced goods. Each correctly solved puzzle is worth GHS 0.20. We opted for a rather difficult task, as it has been demonstrated that such difficult tasks generate

[5] Complete instructions can be found on the author's website (https://sites.google.com/view/kerstin-grosch/research)

higher overestimation (Larrick et al., 2007). Time is limited to 15 seconds per task and the first answer is always logged in as a default to avoid the opportunity of strategically-acting participants who do not work on the task at all, and who are then perfectly able to assess their performance. Afterward, sellers estimate their performance, i.e., the number of solved puzzles, and earn an additional GHS 5 if their guess is exactly correct.

2. Contract offer: Sellers can decide whether they want to offer the buyer a contract and if so how *many* goods they want to offer. Sellers' offers are limited to their individual performance guess. They can earn GHS 1 for each good sold to the buyer. Goods can also be sold for a unit price of GHS 2.5 at an "alternative market" which is basically the experimenter. However, these prices only come into effect when the buyer accepts the seller's offer. Otherwise, goods keep the value of GHS 0.20. We provide sellers with a calculation table to facilitate payoff calculations and to ensure that sellers make informed decisions.

3. Contract acceptance: Buyers, on the other end, receive an endowment of GHS 4. They can decide to invest it in the seller's production to enhance the produce value or to keep the endowment. In the rare event of a seller not offering a contract, the buyer keeps the endowment and the investment game ends here for this dyad. If buyers decide not to invest, the value of the seller's production cannot be "enhanced" and goods are only worth GHS 0.20 in the alternative market.

Buyers can base their decision on the following information. They learn about sellers' performance in the "trial phase" of the real-effort task.[6] They also learn about the number of goods the seller is offering to buy. If the buyer accepts the contract, GHS 4 are invested in the seller's production. The buyer receives GHS 1 for each good the seller sells to the buyer. If the buyer rejects the offer, the buyer keeps the endowment. In this case, the buyer and the seller are informed about their payoffs and the game ends. Payoffs for sellers are as described in Step 2.

4. Information about total output and occurrence of a production shock: After contract conclusion, sellers and buyers are informed about the real total output produced and the corresponding realized payoff. The total output equals the real performance in the real-effort task if no shock affected "the production." A shock affects the production with a 50% chance and output is either topped up by two goods or reduced by two goods (each case is equally likely). The two parties learn wheater a shock has occurred or not and the resulting total output is displayed. In the shock condition, subjects are not informed about the real performance but only about the resulting total output.

[6] The design of the trial phase is similar to the actual investment game. The style of the puzzles is similar as well as the payment, which is GHS 0.20 per solved puzzle. However, in the trial phase, they do not learn about the investment game, yet. Therefore, we do not consider it strictly as an element of the investment game.

5. Selling of goods: To remind parties of the details of the concluded contract, the contract is summarized and displayed. Then, sellers can decide on the amount of goods they ultimately want to sell to the buyer. We provide sellers (again) with a calculation table though this time the payoff expectation and the realized payoff when complying to the contract is displayed at the top of the screen. Sellers then decide how many goods of their total output they want to sell to the buyer and to the alternative market.

In our set-up, sellers breach (over-fulfill) the contract by selling less (more) goods to the buyer than offered and agreed upon with the buyer. Breaching the contract is always monetarily beneficial for the seller. Similar moral-hazard situations exist in real life. For example, in a contract-farming situation: Once a principal has invested in the production, farmers may reduce their effort exerted in the production or only sell low quality produce.

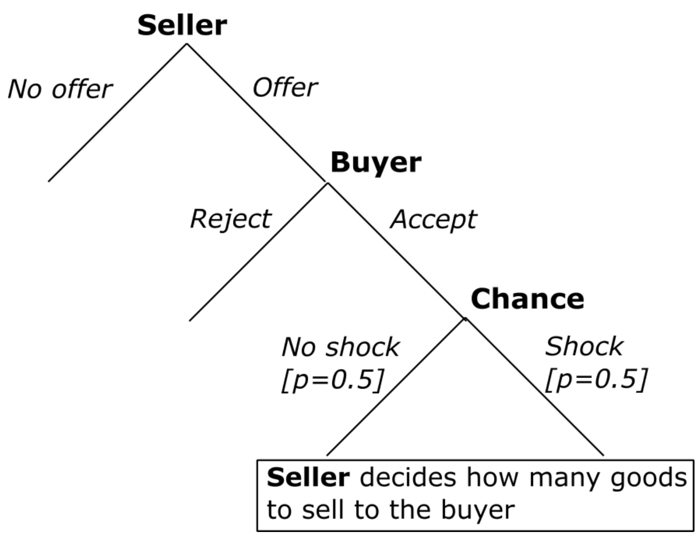

Figure 1: Sequence of actions in the modified investment game

In the previous section we derived propositions for how changes in the expected payoffs ($|\pi_0|$) affect contract breach. In our design, expected payoffs are directly related to the sellers' estimations about their performance (P_{exp}): $\pi_e = Q_{offered} * GHS1 + (P_{exp} - Q_{offered}) * GHS2.5$, whereas π_e is the expected payoff, $Q_{offered}$ is the the seller's offered number of goods, and P_{exp} is the expected performance. Here, we can see that a deviation of P_{exp} from the real performance indicate misjudgments in own performance captured by under-/overestimation of performance. The individual differences in subjective monetary losses and gains correspond to individual differences in under-/overestimation. Hence, we use under-/overestimation of performance as variable of interest in the analysis later on.

Manipulation of overestimation (anchor treatment)

To exogenously manipulate performance guesses to increase payoff expectations, we use an experimental treatment. In the treatment condition, we pose the simple yes-or-no question "Do you think you have produced more than 20 goods?" just before they estimate their performance. The number of 20 goods in the question serves as an anchor. Recall that sellers worked on 25 puzzles in total. Less than 1% of participants actually reached a total performance of over 20 goods. Therefore, we set the anchor extremely high and effectively not reachable for the majority of participants. However, participants use the number 20 to make their subsequent performance guess and start iterating downwards until they reach their final guess (Tversky and Kahneman, 1975). The number 20 might sound arbitrary, but we chose that number for at least two reasons. First, we wanted to implement a substantially high anchor to raise performance guesses, and second, it is an even focal number that mitigated skepticism and avoided that participants contemplate the purpose of the question.

General procedures, elicitation of preferences, and ex-post survey

Before starting a session, we ensured there were an even number of participants in the session to guarantee one-to-one interactions in the investment game. Subjects were fully informed about the entire course of action before they took actual decisions. We explained the game twice, first with a predominantly descriptive text of each step and then using mainly pictures. Control questions were used to make sure that participants understood the sequence of actions.

Prior to the investment game, we elicited (social) preferences. More precisely, we measured social value orientation (Murphy et al., 2011), risk preferences (Eckel and Grossman, 2002), inequality aversion à la Blanco et al. (2011), and loss aversion à la Gächter et al. (2007) which will be used to check that our main variable of interest, *under-/overestimation*, is independent of individual characteristics as well as controls in the analysis. The measure of loss aversion allows us to test whether loss aversion amplifies the effect of overestimation.

The ex-post questionnaire started off with questions about the experiment such as payoff satisfaction and the sellers' feelings when learning about the realized payoff from the contract. We use this survey data to scrutinize the reference-point hypothesis. We also collected data on guilt aversion (Cohen et al., 2011), experience in experiments and on how many other persons the participants knew in the room. At the very end, we asked participants about their socio-demographic background, such as education and income.[7]

[7] As well as that, we play a modified version of the die-rolling game from Fischbacher and Föllmi-Heusi (2013). Since we played it at the very end of the experimental session, it cannot influence any of the experimental data. For brevity, we will therefore not explain the game's details.

Discussion of design decisions

In real life, it is common for an investor to make an offer to a producer. Then, the seller can accept the offer and start producing. In our set-up, sellers (and not buyers) offer the contract. With this procedure, we want to assure a certain level of sellers' commitment to the contract. Commitment is important to test our hypotheses as otherwise sellers would primarily breach the contract because they do not feel obliged to comply to the contract.[8]

For the same reason, we allowed for self-selection into contracts. That is, sellers could decide to offer a contract and buyers to accept the contract. Not allowing for voluntary agreements and assigning predefined contracts could have led to a situation in which sellers breach primarily because they perceive the contract as unfair. This procedure, i.e., allowing subjects to choose their preferred contract (if any), is common practice when studying behavior after contract conclusion (e.g., Fehr et al., 2011; Bartling and Schmidt, 2015).

The price structure was another major design decision. The price in the alternative market is higher than the price that the buyer pays for a good once a contract is concluded. In real life, e.g., contract farming can manifest such a price structure. Since the buyer invests in the seller's production, buyers may pay lower prices to redeem the investment compared to prices offered at the spot market. This creates a moral hazard situation and agents can be tempted to sell to the alternative market.

2.3 Recruitment, sample, and data collection

The sample consists of students recruited from the two largest cities in Ghana, the capital Accra and the university city Kumasi. Our study was publicly announced without any reference to the content of the study and students could register for a particular session in advance. To attract participants, we announced a show-up fee of GHS 20 and used non-monetary incentives such as certificates for participation and a free drink for each participant. Overall, the majority of participants were male (about 70% men and 30% women). On average, participants were 21 years old and said to have a disposable monthly income of GHS 200 (corresponds to about US $40). Most of the students were participating in a lab experiment for the first time.

[8] Let us imagine that the buyer makes the offer. In this case, sellers might accept the offer mainly because the buyer's investment enhances the value of their goods and holds substantial benefits in any case, breaching or complying, compared to a situation in which the seller would reject the contract. However, the seller might perceive the buyer's offer as unfair from the start and, as a consequence, might breach the contract. In such a set-up, the perceived unfairness effect may superimpose an effect from unmet payoff expectations. It would have been a major challenge to capture sellers' fairness perceptions of the contract in a reliable way; if it had not been an impossible endeavor. Therefore, we decided to turn the more intuitive order around and let the seller make the offer.

The data was collected in November and December 2016. In total, we conducted 52 sessions over the course of five weeks at the aforementioned two locations. The sessions comprised between 16 to 24 participants each. We implemented four different treatments. Treatments and shock conditions were randomized within sessions. We analyze data only from two treatments in this paper – from the control treatment explained in section 2.2 and the anchor treatment. In these two treatments, a total of 478 participants took part, that is, 239 sellers and 239 buyers. We received an ethical approval for this study from the ethical board of social sciences at the University of Goettingen.

3 Data and results

3.1 Descriptive statistics

In Table 1, we summarize observations and differentiate between sellers who correctly guess their performance, and underconfident and overconfident sellers. Sellers categorized as overconfident made a guess that exceeded their real performance, participants with a performance guess below their real performance are categorized as underconfident, and participants with a correct guess evaluated their performance spot on.

Contract breach (*breach*) is defined as the difference between the number of goods promised and eventually sold to the buyer. In case that the promised number of goods to the buyer exceeds the total output of goods produced (40 sellers in our sample), contract breach is defined as the total number of goods sold to the alternative market. This definition of contract breach allows us to focus on deliberate contract breach.

Table 1: Estimation types and compliant/non-compliant behavior

	correct guess		underconfident		overconfident		Total	
compliance	7	(50.00%)	23	(54.76%)	52	(44.83%)	82	(47.67%)
breach	7	(50.00%)	5	(11.09%)	63	(54.31%)	75	(43.60%)
over-fulfillment	0	(00.00%)	14	(33.33%)	1	(1.32%)	15	(11.36%)
total	14	(8.14%)	42	(24.42%)	116	(67.44%)	172	(100.00%)

In total, 186 contracts were concluded. Due to power outages in the field, we lost certain variables for 14 observations leaving us with 172 observations to analyze. Observation numbers in different stages and treatments can be looked up in Table 7 in the appendix.

The majority of sellers, 67.44%, overestimate their performance in the real-effort task whereas 8.14% of sellers correctly assess, and 24.42% underestimate their performance. A histogram

of under-/overestimation in our sample can be found in the appendix in Figure 3. Generally, we see that a large fraction of sellers (47.67%) comply with their contracts whereas 43.60% breach the contract and a small fraction of 11.36% over-fulfill the contract (*breach*< 0).
The average performance in the real-effort task is 11.84 and does significantly differ between underconfident (12.90) and overconfident (11.34) sellers (Mann-Whitney test, $p = 0.012$), but not between sellers who correctly guessed their performance (12.86), and over- or underconfident sellers (Mann-Whitney tests, both $p > 0.158$). Overconfident sellers are off the mark by an average of 5.19 goods. Contract breach among overconfident agents is 2.28 on average, sellers who correctly guessed their performance shortchange their buyers by an average of 2.57 goods, and underconfident sellers over-fulfill the contract by 0.55 units on average. The difference in absolute breaching between under–and overconfident sellers, as well as between underconfident sellers and correct estimators is statistically significant (Mann-Whitney tests, all $p < 0.002$). Also, overconfident sellers, as well as correct estimators, breach the contract significantly *more often* than underconfident sellers (*overconfident sellers:* $\chi^2(1) = 41.7093$, $p < 0.001$, *correct estimators:* $\chi^2(1) = 9.0505$, $p = 0.003$). The difference between the breaching level of overconfident sellers compared with agents who correctly guessed their performance is not statistically significant (Mann-Whitney test, $p = 0.828$).
On average sellers offered 63.33% of their performance guess to the buyer and this share does not differ across estimation types. This indicates that there are no differences in prudence or strategic behavior when making a certain offer across different types. However, different types end up with varying actual shares of output that have to be sold to the buyer to comply with the contract: While underconfident sellers have to sell on average only 44.38% (5.86 goods) of their total output, correct estimators have to sell on average 65.56% (8.36 goods), and overconfident sellers are supposed to sell on average 79.91% (8.74 goods) of their produce to the buyer. Thus, we control for the offers in the analysis.
Sellers may stick to the particular share they offered but not to the absolute amount promised in the contract. However, this is not confirmed by the data. While there were no statistical differences in the initial relative offers across estimation types (Mann-Whitney test, $p = 0.468$), the final shares of the produce sold to the buyer differ significantly between over- (59.40%) and underconfident (48.69%) sellers (Mann-Whitney test, $p = 0.015$), suggesting that the contracts indeed created a commitment.

Result I: *Overall, we find that a share of 48% of sellers complied with contracts. Overconfident sellers and correct estimators breach contracts more often and to a higher extent than underconfident sellers.*

In the following, we will check whether individual overestimation is independent of individual characteristics and preferences to avoid misinterpretations of regression results later on. In Table 2, we present results of OLS regressions from different characteristics on *under-/overestimation*; standard errors in parentheses. Only one regressor is included at a time. The variable *under-/overestimation* is a continuous measure and is calculated by the differ-

Table 2: Test of independence for overestimation

	age	income	share offered	loss aversion	risk seeking	SVO angle
Under-/overest.	-0.257	7.736	-0.009	-0.001	0.001	0.288
	(0.033)	(4.106)	(0.017)	(0.033)	(0.029)	(0.258)
Constant	21.387	245.07	3.784	3.857	3.125	26.51
	(0.174)	(22.066)	(0.917)	(0.177)	(0.157)	(1.386)
Observations	172	172	172	172	172	172

Standard errors in parentheses

ence between performance guess and real performance. The variable is zero for sellers who correctly guessed their performance, positive for overconfident and negative for underconfident sellers. We see that none of the coefficients are statistically significant, implying that different levels of *under-/overestimation* are independent of social preferences such as Social Value Orientation (*SVO angle*) or inequality aversion (*share offered* in an ultimatum game). The literature suggests that men might be more overconfident than women (e.g., Barber and Odean, 2001). We do not find a gender difference in *under-/overestimation* in our sample (Mann-Whitney tests, $p = 0.237$). Moreover, students enrolled in an economics program are no different than students enrolled in other programs with respect to *under-/overestimation* (Mann-Whitney test, $p = 0.963$). In the next section, we test our propositions from section 2.1 with regression analyses.

3.2 Estimation results

Does under-/overestimation affect contract breach?

In the following, we graphically resume the relationship between overestimation and contract breach. In Figure 2, underconfident sellers range left from *0* whereas overconfident sellers range right from *0* on the x-axis. Sellers with a guess spot-on can be found on the y-axis. The upward trend of the fitted line suggests that with rising under-/overestimation, contract breach increases.

Balancing tests in Table 2 already indicated that there are no systematic differences in

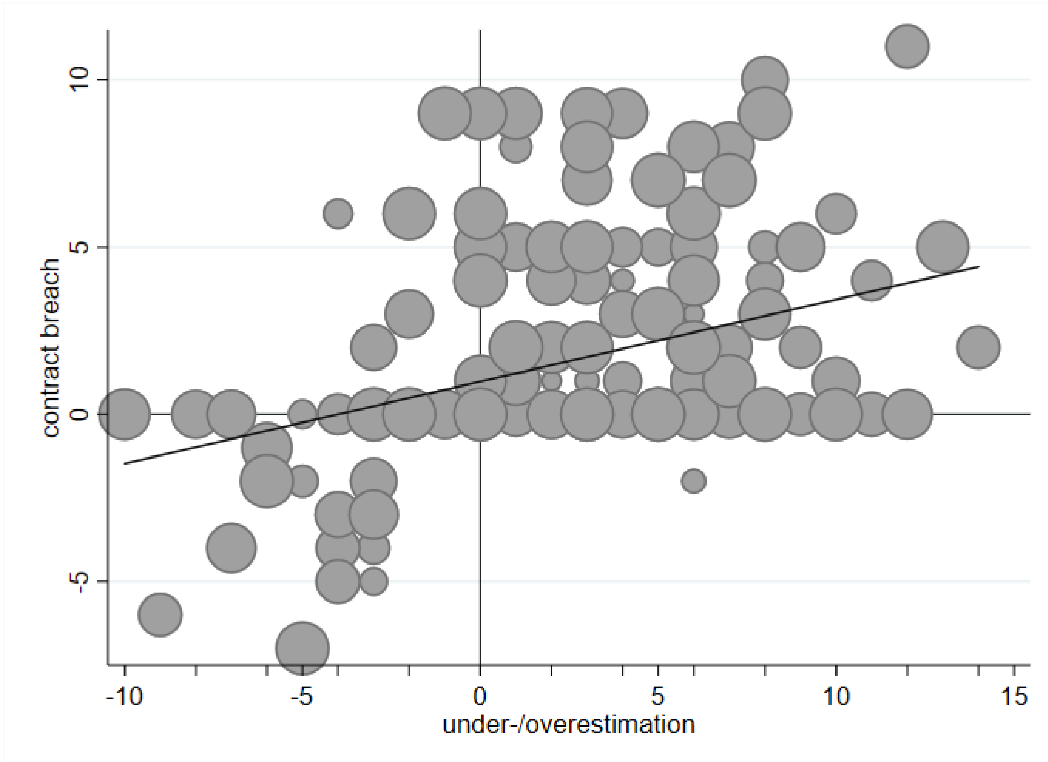

Figure 2: Effect of overestimation on contract breach (N=172)

covariates across estimation types. Additionally, we use an instrumental variable approach by using our anchor treatment as an instrument for over-/underestimation to further test whether our measure of under-/overestimation is exogenous.

The underlying OLS model we use is of the following form:

$$Y_i = \beta_0 + \beta_1 Q_{under-/overestimation} + \beta_2 S_i + \beta_3 P_i + \beta_4 C_i + \epsilon_i,$$

where Y_i is the level of contract breach and measured as defined in section 3.1 (*breach*). Note, that Y_i is *0* when sellers comply, becomes positive when sellers breach, and negative when sellers over-fulfill the contract. $Q_{under-/overestimation}$ represents the level of under-/overestimation. S_i is an indicator variable which is positive if a shock affected sellers' output and *0* otherwise (*shock*); P_i is sellers' performance (*performance*)[9]; C_i represents a vector of covariates at the individual level; ϵ_i is the error term.

In Table 3, we present results from different specifications of the OLS regressions in models (1) and (3). In a second step, we estimate the regression model using an instrumental variable

[9]We control for performance in all our regressions since sellers with very high performances cannot overestimate themselves as strongly as sellers with very low performances.

approach. In the IV approach, we use a dummy for the anchor treatment as an instrument for under-/overestimation and re-estimate the model specifications (1) and (3). The results are presented in models (2) and (4). Results from all models demonstrate that with rising *under-/overestimation* contract breach significantly increases. The IV regressions (models (2) and (4)) reveal that the reported effects from overestimation are causal.[10] Results from the first stage regression demonstrate highly significant effects from the anchor treatment on overestimation and can be looked up in the appendix in Table 8.[11]

Table 3: Effect of overestimation on contract breach (OLS and IV results)

VARIABLES	(1) OLS	(2) IV	(3) OLS	(4) IV
Over-/underestimation	0.291***	0.433***	0.294***	0.377***
	(0.0494)	(0.166)	(0.0510)	(0.142)
Shock	0.560	0.575	0.606	0.620
	(0.453)	(0.462)	(0.471)	(0.456)
Performance	0.199***	0.261***	0.165**	0.207**
	(0.0607)	(0.0862)	(0.0716)	(0.0884)
Constant	-1.819**	-2.937**	-2.999	-3.687
	(0.809)	(1.411)	(2.978)	(3.186)
Observations	172	172	172	172
Adjusted R-squared	0.165	0.124	0.193	0.179
Covariates[a]	No	No	Yes	Yes

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[a]Covariates: age, income, share offered in ultimatum game, loss aversion, risk-seeking, SVO angle, guilt aversion, dummy for female, dummy for experiment's location (Kumasi or Accra), income per month (in real life), age in years, dummy for enrollment in an econ program.

Performance has a significant positive effect. A reason can be that the opportunity costs of compliance rise with increasing performance. Therefore, sellers with a high performance might be more prone to breach the contract to secure a high payoff from the lucrative alternative market. Covariates are included in models (3) and (4). The only covariates that are significant at a 5%-level are *location*, i.e., subjects from the university in Kumasi breach

[10]The Wu-Hausman test (tested for non-robust standard errors) reveals that we cannot reject the hypothesis that the variable *under-/overestimation* is exogenous ($p > 0.374$, models 2 and 4).

[11]On average, sellers' guesses are 3.1 units higher under the anchor treatment compared to the no-anchor treatment. This difference is significant (Kolmogorov-Smirnov test $p = 0.002$). The anchor did not alter the distribution of overestimation. To prove that, we reduce individual overestimation under the anchor by the mean of 2.201. When comparing the adjusted anchor distribution with distribution of overestimation in the control treatment, there is no significant difference (Kolmogorov-Smirnov test, $p = 0.311$). This demonstrates that the anchor increases guesses to a similar extent at all different performance levels.

less than subjects from the university in Accra (Model (4), $p = 0.022$), and the female dummy which is also negative indiciating that women may breach less than men (Model (4), $p = 0.045$). In these models, the effects from overestimation remain similar compared to the models without controlling for covariates.

Result II: *With increasing under-/overestimation, sellers breach the contract to a higher extent. The relationship between under-/overestimation and contract breach is positive and causal.*

Who breaches and who over-fulfills the contract?

In the following, we distinguish between sellers who breached and sellers who over-fulfilled the contract. We compare these two groups with the subjects who complied with the contract. Results of probit model specifications are reported in Table 4. Models 1 and 2 focus on the analysis of sellers who breached whereas models 3 and 4 focus on sellers who over-fulfilled the contract.

For the following analysis, we split up the variable *under-/overestimation* into two variables, *overestimation* and *underestimation*. The variable *overestimation* is 0 for subjects underestimating or correctly guessing their performance and increases continuously with the level of overestimation. Similarly, the variable *underestimation* is 0 for subjects overestimating or correctly guessing their performance and increases continuously with the level of underestimation. It is striking that *underestimation* is the variable that is highly significant. More precisely, underconfident sellers are more likely to over-fulfill contracts (models 3 and 4) and less likely to breach contracts (models 1 and 2). *Overestimation* is positive for contract breach but not statistically significant (models 1 and 2). This indicates that *underestimation* rather than *overestimation* is a predictor for breaching or over-fulfilling respectively. Hence, the results are only partly in line with Proposition I.

Result III: *With increasing underestimation, sellers are more likely to over-fulfill the contract and vice versa less likely to breach the contract. We do not find significant effects of overestimation on the probability of breaching or over-fulfilling the contract.*

What determines the level of contract breach?

To determine if over- and underestimation affect the particular level of contract breach, we reduce the sample to the observations in which agents breach the contract (N=75). We

Table 4: Probit regression models on contract breach and contract over-fulfillment

VARIABLES	(1) breach	(2) breach	(3) over-fulfill	(4) over-fulfill
Overestimation	0.0549	0.0565	-0.0932	-0.226
	(0.0353)	(0.0372)	(0.0987)	(0.159)
Underestimation	-0.236**	-0.255**	0.240***	0.160*
	(0.112)	(0.111)	(0.0748)	(0.0961)
Performance	0.0449	0.0380	0.0235	0.120
	(0.0320)	(0.0354)	(0.0578)	(0.0906)
Shock	0.411*	0.395*	0.0163	-0.272
	(0.212)	(0.225)	(0.372)	(0.455)
Constant	-0.931*	-1.144	-1.735**	-0.320
	(0.492)	(1.550)	(0.828)	(3.499)
Observations	157	157	97	97
Covariates[a]	No	Yes	No	Yes
Pseudo R-squared	0.0812	0.144	0.299	0.417

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

[a]Covariates: age, income, share offered in ultimatum game, loss aversion, risk-seeking, SVO angle, guilt aversion, dummy for female, dummy for experiment's location (Kumasi or Accra), income per month (in real life), age in years, dummy for enrollment in an econ program.

expect that particularly overconfident sellers with a pronounced individual loss aversion breach contracts to a higher extent than less overconfident sellers with a less pronounced individual loss aversion. To test this Proposition IIa, we include an interaction term between loss aversion (λ_i) and overestimation ($Q_{overest.}$) in an OLS regression model.

$$Y_i = \beta_0 + \beta_1 Q_{overest.} + \beta_2 \lambda_i + \beta_3 \lambda_i Q_{overest.} + \beta_4 Q_{underest.} + \beta_5 P_i + \beta_6 S_i + \beta_7 C_i + \epsilon_i,$$

with all variables defined as previously.

Results are reported in models (3) and (4) in Table 5. We see that the interaction between *overestimation* and *loss aversion* is positive and significant on a 5%-level. In models (5) and (6), we demonstrate that these results remain robust when the number of sellers' offered goods are included. Focusing on models (1) and (2), we cannot confirm Proposition IIb that with rising underestimation, sellers who breach, breach less.

Result IV: *With increasing individual loss aversion, overconfident sellers breach the contract to a higher extent.*

Table 5: Effect of overestimation and loss aversion on breaching level (OLS regression models)

VARIABLES	(1) breach	(2) breach	(3) breach	(4) breach	(5) breach	(6) breach
Overestimation (Overest.)	0.121	0.0996	-0.241	-0.314	-0.265	-0.373*
	(0.0987)	(0.108)	(0.186)	(0.200)	(0.184)	(0.199)
Loss aversion (loss av.)	0.168	0.226	-0.220	-0.190	-0.172	-0.133
	(0.164)	(0.182)	(0.233)	(0.246)	(0.232)	(0.243)
Underestimation	0.0371	-0.0149	0.0379	-0.0188	0.0814	0.0642
	(0.474)	(0.525)	(0.461)	(0.505)	(0.455)	(0.497)
Overest. × loss av.			0.0915**	0.103**	0.0782*	0.0945**
			(0.0403)	(0.0429)	(0.0405)	(0.0423)
Performance	0.289***	0.278**	0.271***	0.249**	0.225**	0.175
	(0.0935)	(0.106)	(0.0912)	(0.103)	(0.0939)	(0.108)
Shock	0.151	-0.0979	0.0761	-0.217	-0.0463	-0.395
	(0.650)	(0.734)	(0.632)	(0.708)	(0.627)	(0.700)
Seller's offered goods					0.163*	0.204*
					(0.0960)	(0.111)
Constant	-0.321	-3.496	1.327	-2.798	0.450	-5.032
	(1.620)	(4.463)	(1.733)	(4.304)	(1.786)	(4.389)
Observations	75	75	75	75	75	75
R-squared	0.141	0.190	0.201	0.263	0.234	0.303
Covariates[a]	No	Yes	No	Yes	No	Yes

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[a]Covariates: age, income, share offered in ultimatum game, risk seeking, SVO angle, guilt aversion, dummy for female, dummy for experiment's location (Kumasi or Accra), income per month (in real life), age in years, dummy for enrollment in an econ program.

Does the occurrence of shocks facilitate contract breach?

We investigate how the occurrence of a shock affects contract breach (Proposition III) in the following. In general, when focusing on the frequency of breaching, 65.33% of the sellers breach the contract in the shock condition (non-deterministic environment) whereas only 34.67% of sellers engage in breaching in the no-shock condition (deterministic environment). This means that about twice as many sellers breach in the shock condition compared to the no shock condition. This is confirmed by regression analysis as seen in models (1) and (2) in Table 4 when it is about the decision to breach or not to breach. The non-significant results of the dummy variable for the shock condition for the sellers who breach in Table 5 suggest that the shock does not affect the level of breaching.

Result V: *The occurrence of shocks increases the likelihood of breaching. However, it does* *not affect the level of breaching contracts.*

3.3 Evidence of reference-dependent behavior

In this section, we will discuss evidence of expectation-based, reference-dependent preferences in our experiment. For this purpose, we will first recall the results presented in the previous section. Moreover, we examine data from the ex-post questionnaire on feelings when learning about the status quo and sellers' satisfaction with the payoff from the contract. All test results presented are based on data used in the previous analysis (N=172).

Estimation results

We find an effect of overestimation on contract breach conditional on loss aversion. This effect is robust to controlling for the individual offers made. The perception of a loss can only eventuate if sellers indeed compare the status quo with the reference payoff. This result clearly demonstrates that a negative deviation from payoff expectations, the alleged reference point, is perceived as a loss.

Income satisfaction

When sellers compared their realized payoff with the alleged reference point, they might be less satisfied if they failed to reach it compared to a situation where they managed to reach it. In other words, the higher the experienced subjective loss, the less satisfied sellers might be with their actual payoff.

In the following, we test the effect from overestimation on payoff satisfaction stated in the ex-post questionnaire of the experimental session. We asked sellers "How satisfied are you with your final payoff from the production phase?" They could answer on a scale from 1 to 10 where 1 meant very low satisfaction and 10 very high satisfaction. Results are presented in Table 6. The variable *payoff* entails the final payoff from the investment game. In both models, we control for several covariates such as sex, location, guilt aversion, social value orientation, age, and enrollment in an economics program. Generally, economic students are significantly less satisfied with their payoff from the investment game and guilt aversion is weakly positively correlated with income satisfaction. The other covariates, that are not presented in the table, are not significant.

To examine whether overestimation reduces payoff satisfaction for a given payoff level, we include an interaction term of *overestimation* and *payoff.* The interaction effect is negative and highly significant. This suggests that indeed overconfident sellers are less satisfied with

Table 6: Seller's payoff satisfaction

	Income satisfaction	
	(1)	(2)
Overestimation	0.112	0.146
	(0.197)	(0.112)
Underestimation	-0.093	0.163
	(0.070)	(0.356)
Payoff	0.026	0.036
	(0.032)	(0.033)
Overestimation × *Payoff*	-0.012**	-0.014**
	(0.005)	(0.005)
Underestimation × *Payoff*		-0.012
		(0.014)
Performance	0.052	0.054
	(0.072)	(0.072)
Shock	0.079	0.089
	(0.290)	(0.289)
Constant	2.888	2.676
	(1.848)	(1.821)
Observations	172	172
R-squared	0.168	0.171
Covariates[a]	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[a]Covariates: age, income, share offered in ultimatum game, loss aversion, risk-seeking, SVO angle, guilt aversion, dummy for female, dummy for experiment's location (Kumasi or Accra), income per month (in real life), age in years, dummy for enrollment in an econ program.

their payoff. The results are robust when we include an interaction term of *underestimation* and *payoff* in model (2), the effects reported above are robust and *breach (absolute)* has no significant effect on payoff satisfaction.

Disappointment

In the ex-post questionnaire, we asked sellers how they felt after they learned how much they produced and we specified the following four answers "I felt disappointed," "I felt joyful," "I was surprised," and "something else." We would expect that overconfident sellers would feel more disappointed than sellers who were underconfident or correctly guessed their performance. And this is exactly what we find. From 116 overconfident sellers, 43.1% stated that they were disappointed, 28.45% were surprised and 12.93% joyful. In contrast, only three subjects (7.14%) of the 42 underconfident sellers felt disappointed, 45.24% surprised, and 33.33% joyful. Overall, about 16% stated that they felt "something else" than our preset answers. The distribution of stated feelings is significantly different between sellers who correctly estimated their performance/underestimated their performance and overconfident sellers ($\chi^2(1) = 21.914,\ p < 0.022$).

We indeed find that disappointed sellers shortchange, on average, 1.5 goods more compared to non-disappointed sellers. This difference is statistically significant using a Mann-Whitney test ($p = 0.002$).[12] This result could be a mere correlation rather than a causal relationship between biased expectations and contract breach. However, it may be another hint that sellers actually compare their realized payoff with their expected payoff and discharge the negative emotions arising from the comparison by breaching (à la Loewenstein and Lerner, 2003).

4 Discussion

Reference payoffs and why people may respond to it

There are different reasons why and when people respond to reference payoffs. Payoff expectations might simply serve as a goal that agents want to achieve (Oettinger, 1999; Fehr and Goette, 2007). In contract situations, it has been found that agents' feelings of entitlement to a reference payoff plays a role. When principals withdraw part of the agent's reference payoff, agents feel betrayed and, consequently, breach the contract (Fehr et al., 2011). This should not play a dominant role in our context because agents make the contract offer and have the opportunity to betray the buyer, and not the other way around.

Another reason for responding to a reference payoff may be negative emotions such as disappointment or dissatisfaction (e.g. Bell, 1985; Loomes and Sugden, 1986). Breaching the contract might serve as an "emotion-relieving action" that can counteract the source of disappointment (Loewenstein and Lerner, 2003).

The discussed motives for responding to reference payoffs may be intertwined. Depending on the context of the situation and depending on individual characteristics and preferences one or the other motive may play a more prominent role. Our results, particularly those of section 3.3, support that emotions such as disappointment and dissatisfaction may explain

[12] To test whether disappointed sellers breach more, we created a dummy variable that takes the value 1 if sellers stated that they were disappointed and takes the value *0* if seller stated "joyful," "surprised" or "something else." We tested whether the means of absolute breach was different between the two groups with a Mann-Whitney test.

at least part of behaving non-compliantly. In addition, emotions such as an elated feeling à la Bell (1985) and Loomes and Sugden (1986) may be part of the explanation of the positive correlation of underestimation and over-fulfillment of contracts. The unexpected gain might feel like a windfall of money and may make people more willing to share their income (Kameda et al., 2002; Carlsson et al., 2013).

Production shocks and possible underestimation of a self-delusion effect

Apart from the results on payoff expectations in general, our experimental data implies that the occurrence of production shocks increases the likelihood of contract breach. The shock might serve as a convenient justification for breaching the contract since here sellers can shift accountability for the failure in performance judgment to the random event. Accordingly, Blanton et al. (2001) found that people are concerned about their self-image and want to avoid cognitive dissonance. Therefore, the cognitive bias from overconfidence might be self-sustaining via a confirmation bias or the "ego-utility" people receive from having a high estimation of themselves (Köszegi, 2006). Real life is full of non-predictable events. The farmer might complain about the weather, the manager of a factory about soaring supplier prices, and the head of a service center about unreliable staff. The examples illustrate that it is hard to imagine a situation where there is no opportunity to delude oneself and shift accountability of own failure to another event or persons. In our experiment, we may underestimate the effect of self-delusion. The shock can serve as an excuse but many participants were taking part in an experiment for the first time, and also we conducted the experiments with tablets and, therefore, they may have felt observed. Hence, in our set-up, participants may not have responded as strongly to the wiggle room in the shock condition as they may have in a real-life situation.

5 Conclusion

In contract situations, overestimating one's own performance can lead to inflated payoff expectations. If overconfident agents use payoff expectations as a reference point, they might sense a subjective loss as soon as they realize that their performance falls short of this reference payoff. As a consequence, agents may try to balance off these losses by breaching contracts. In this study, we find evidence of this linkage. Results from our lab experiment show that overconfident agents breach contracts to a higher extent with increasing loss aversion. By manipulating overestimation in a treatment with a performance anchor, we can show that the effect from under-/overestimation is causal. Moreover, there are several indications that this effect can be explained by reference-dependent preferences. These indicators reveal that agents compared their status quo (realized payoff from the contract) with the alleged reference point – their payoff expectations. First, we find evidence that higher levels of loss aversion stimulate contract breach for overconfident agents. Second, our post-experimental questionnaire reveals that agents who overestimated their payoff expectations are more often disappointed in the moment of learning about their status quo than agents who underestimated their payoff expectations. Third, we find that agents become less satisfied with their payoff from the contract with rising overestimation.

Surprisingly, the effect of underestimation is somewhat stronger than the effect of overestimation. Underconfident agents breached significantly less than overconfident ones and the magnitude and significance of underestimation on contract breach is more pronounced than the effect of overestimation. Moreover, underconfident agents over-fulfill contracts substantially more often and to a higher extent than agents who correctly assessed or overestimated their performance.

A novelty of our study is that we show how a common heuristic such as overconfidence may lead to non-compliant behavior in contract settings by forming inflated payoff expectations. In our set-up, the reference point forms endogenously based on one's own performance misjudgment as compared to, for instance, Baillon et al. (2019) who examine the share of people orientating on different types of reference points (e.g. status quo, expectation-based) and where the reference point is stochastic. Furthermore, in contrast to the majority of studies that focus on ex-ante behavioral change under reference points (e.g., Camerer et al., 1997; Rabin, 2000; Abeler et al., 2011), we focus on an ex-post impact of unmet expectations. Moreover, this study contributes to the scarce empirical and experimental literature on reference-dependent preferences. Particularly, we have investigated the theoretical idea that unmet expectations lead to disappointment and a utility loss, the more pronounced the individual loss aversion (e.g., Bell, 1985; Loomes and Sugden, 1986; Gul, 1991; Koeszegi and Rabin, 2006). In line with this theory, we find that individual loss aversion increases contract breach when the status quo falls short of the reference payoff.

Besides the literature on reference-dependent preferences, we draw on and contribute to the literature on performance misjudgment and especially overconfidence. The majority of the literature examines implications of overconfidence for financial markets and managerial decision-making (e.g. Barber and Odean, 2001; Scheinkman and Xiong, 2003). It has been shown that overconfidence can lead to deadweight losses in markets. Since overconfident agents are more prone to overestimate future market shares, they enter these markets excessively (Camerer and Lovallo, 1999; Koellinger et al., 2007). Overconfidence can also affect contract outcomes. Overconfident job seekers may underestimate the risks involved in certain contracts, misjudge effort that has to be exerted, and any prospective wage that will be gained. This way, contract and incentive designs are adjusted on the employer's side, and sorting decisions and effort provision are affected on the employee's side by overconfidence, potentially leading to non-optimal outcomes (De la Rosa, 2011; Sautmann, 2013; Larkin and Leider, 2012; Santos-Pinto, 2012; Hoffman and Burks, 2020).

We are careful when making strong inferences for policymakers since we used a stylized lab experiment and, hence, external validity may be constrained. Still, our results suggest that supporting people in judging their own abilities correctly, especially those who are highly loss averse, can reduce contract breach. It has been shown that overconfidence decreases very slowly over time, if at all, even when overconfident agents gain experiences in certain tasks (Hoffman and Burks, 2020). In a different study, we will examine to what extent feedback mechanisms can enhance performance judgment and potentially curb contract breach (Fischer and Grosch, 2020). Furthermore, policies should ensure that principals do not raise unrealistic expectations to avoid spuring agents' overconfidence and high payoff expectations. To curb the temptation of self-delusion in non-deterministic environments, bookkeeping and documenting meaningful events at daily work could support agents' contract compliance. However, in some instances, being overconfident may even affect outcomes positively and increase real performance after all by reducing emotions such as anxiety of failure or by actions striving to avoid an anticipated loss (Compte and Postlewaite, 2004). Our experimental design does not allow the investigation of these effects. We leave that open for future research.

References

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011). Reference points and effort provision. *The American Economic Review 101*(2), 470–492.

Alm, J., G. H. McClelland, and W. D. Schulze (1992). Why do people pay taxes? *Journal of Public Economics 48*(1), 21–38.

Baillon, A., H. Bleichrodt, and V. Spinu (2019). Searching for the reference point. *Management Science*.

Barber, B. M. and T. Odean (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics 116*(1), 261–292.

Bartling, B. and K. M. Schmidt (2015). Reference points, social norms, and fairness in contract renegotiations. *Journal of the European Economic Association 13*(1), 98–129.

Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations research 33*(1), 1–27.

Bellemare, M. F. and J. R. Bloem (2018). Does contract farming improve welfare? a review. *World Development 112*, 259–271.

Benoît, J.-P., J. Dubra, and D. A. Moore (2015). Does the better-than-average effect show that people are overconfident?: Two experiments. *Journal of the European Economic Association 13*(2), 293–329.

Blanco, M., D. Engelmann, and H. T. Normann (2011). A within-subject analysis of other-regarding preferences. *Games and Economic Behavior 72*(2), 321–338.

Blanton, H., B. W. Pelham, T. DeHart, and M. Carvallo (2001). Overconfidence as dissonance reduction. *Journal of Experimental Social Psychology 37*(5), 373–385.

Camerer, C., L. Babcock, G. Loewenstein, and R. Thaler (1997). Labor supply of new york city cabdrivers: One day at a time. *The Quarterly Journal of Economics 112*(2), 407–441.

Camerer, C. and D. Lovallo (1999). Overconfidence and excess entry: An experimental approach. *The American Economic Review 89*(1), 306–318.

Carlsson, F., H. He, and P. Martinsson (2013). Easy come, easy go. *Experimental Economics 16*(2), 190–207.

Charness, G. and M. Dufwenberg (2006). Promises and partnership. *Econometrica 74*(6), 1579–1601.

Chemin, M. (2012). Does court speed shape economic activity? evidence from a court reform in india. *Journal of Law, Economics, and Organization 28*(3), 460–485.

Chen, D. L., M. Schonger, and C. Wickens (2016). otree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Cohen, T. R., S. T. Wolf, A. T. Panter, and C. A. Insko (2011). Introducing the gasp scale: a new measure of guilt and shame proneness. *Journal of Personality and Social Psychology 100*(5), 947.

Compte, O. and A. Postlewaite (2004). Confidence-enhanced performance. *American Economic Review 94*(5), 1536–1557.

Dana, J., R. A. Weber, and J. X. Kuang (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory 33*(1), 67–80.

De la Rosa, L. E. (2011). Overconfidence and moral hazard. *Games and Economic Behavior 73*(2), 429–451.

Dixit, A. (2003). Trade expansion and contract enforcement. *Journal of Political Economy 111*(6), 1293–1317.

Eckel, C. C. and P. J. Grossman (2002). Sex differences and statistical stereotyping in attitudes toward financial risk. *Evolution and Human Behavior 23*(4), 281–295.

FAO (2019). Ghana at a glance. *http://www.fao.org/ghana/fao-in-ghana/ghana-at-a-glance/en/*, accessed at 2019–08–12.

Fehr, E. and U. Fischbacher (2004). Third-party punishment and social norms. *Evolution and Human Behavior 25*(2), 63–87.

Fehr, E., U. Fischbacher, and S. Gächter (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human Nature 13*(1), 1–25.

Fehr, E. and L. Goette (2007). Do workers work more if wages are high? evidence from a randomized field experiment. *The American Economic Review 97*(1), 5.

Fehr, E., O. Hart, and C. A. Zehndera (2011). Contracts as reference points. experimental evidence. *The American Economic Review 101*(2), 493–525.

Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in disguise - an experimental study on cheating. *Journal of the European Economic Association 11*(3), 525–547.

Fischer, S. and K. Grosch (2020). Belief update: Can feedbacks counteract contract breach? *mimeo*.

Gächter, S., E. J. Johnson, and A. Herrmann (2007). Individual-level loss aversion in riskless and risky choices. *IZA Discussion Paper No. 2961*.

Garbarino, E., R. Slonim, and M. C. Villeval (2019). Loss aversion and lying behavior. *Journal of Economic Behavior & Organization 158*, 379–393.

Gill, D. and V. Prowse (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review 102*(1), 469–503.

Grossman, Z. and D. Owens (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior & Organization 84*(2), 510–524.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica: Journal of the Econometric Society*, 667–686.

Hart, O. and J. Moore (2008). Contracts as reference points. *The Quarterly Journal of Economics 123*(1), 1–48.

Haushofer, J. and E. Fehr (2014). On the psychology of poverty. *Science 344*(6186), 862–867.

Henrich, J., S. J. Heine, and A. Norenzayan (2010). Most people are not weird. *Nature 466*(7302), 29–29.

Hill, R. V., E. Maruyama, and A. Viceisza (2012). Breaking the norm: An empirical investigation into the unraveling of good behavior. *Journal of Development Economics 99*(1), 150–162.

Hoffman, M. and S. V. Burks (2020). Worker overconfidence: Field evidence and implications for employee turnover and firm profits. *Quantitative Economics 11*(1), 315–348.

Huang, G. and H. Liu (2019). Expectation-based reference-dependent preferences: Evidence from the used-car retail market. *Available at SSRN 2902336*.

Jones, D. (2010). A weird view of human nature skews psychologists' studies. *Science 328*(5986), 1627–1627.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263–291.

Kameda, T., M. Takezawa, R. S. Tindale, and C. M. Smith (2002). Social sharing and risk reduction: Exploring a computational algorithm for the psychology of windfall gains. *Evolution and Human Behavior 23*(1), 11–33.

Klein, B. (2000). The role of incomplete contracts in self-enforcing relationships. *Revue d'économie industrielle 92*(1), 67–80.

Klein, B., R. G. Crawford, and A. A. Alchian (1978). Vertical integration, appropriable rents, and the competitive contracting process. *The Journal of Law and Economics 21*(2), 297–326.

Koellinger, P., M. Minniti, and C. Schade (2007). "i think i can, i think i can": Overconfidence and entrepreneurial behavior. *Journal of Economic Psychology 28*(4), 502–527.

Koeszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 1133–1165.

Köszegi, B. (2006). Ego utility, overconfidence, and task choice. *Journal of the European Economic Association 4*(4), 673–707.

Koszegi, B. (2014). Behavioral contract theory. *Journal of Economic Literature 52*(4), 1075–1118.

Kruger, J. and D. Dunning (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology 77*(6), 1121.

Larkin, I. and S. Leider (2012). Incentive schemes, sorting, and behavioral biases of employees: Experimental evidence. *American Economic Journal: Microeconomics 4*(2), 184–214.

Larrick, R. P., K. A. Burson, and J. B. Soll (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes 102*(1), 76–94.

Loewenstein, G. and J. S. Lerner (2003). The role of affect in decision making. *Handbook of Affective Science 619*(642), 3.

Loomes, G. and R. Sugden (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies 53*(2), 271–282.

Marzilli Ericson, K. M. and A. Fuster (2011). Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *The Quarterly Journal of Economics 126*(4), 1879–1907.

Mazar, N., O. Amir, and D. Ariely (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research 45*(6), 633–644.

Moore, D. A. and P. J. Healy (2008). The trouble with overconfidence. *Psychological Review 115*(2), 502.

Murphy, R. O., K. A. Ackermann, and M. Handgraaf (2011). Measuring social value orientation. *Judgment and Decision Making 6*(8), 771–781.

Oettinger, G. S. (1999). An empirical analysis of the daily labor supply of stadium venors. *Journal of Political Economy 107*(2), 360–392.

Pope, D. G. and M. E. Schweitzer (2011). Is tiger woods loss averse? persistent bias in the face of experience, competition, and high stakes. *American Economic Review 101*(1), 129–57.

Post, T., M. J. Van den Assem, G. Baltussen, and R. H. Thaler (2008). Deal or no deal? decision making under risk in a large-payoff game show. *American Economic Review 98*(1), 38–71.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

Reardon, T., C. B. Barrett, J. A. Berdegué, and J. F. Swinnen (2009). Agrifood industry transformation and small farmers in developing countries. *World Development 37*(11), 1717–1727.

Sackeim, H. A. and R. C. Gur (1979). Self-deception, other-deception, and self-reported psychopathology. *Journal of Consulting and Clinical Psychology 47*(1), 213.

Santos-Pinto, L. (2012). Labor market signaling and self-confidence: Wage compression and the gender pay gap. *Journal of Labor Economics 30*(4), 873–914.

Sautmann, A. (2013). Contracts for agents with biased beliefs: Some theory and an experiment. *American Economic Journal: Microeconomics 5*(3), 124–56.

Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *Journal of Political Economy 111*(6), 1183–1220.

Tanaka, T., C. F. Camerer, and Q. Nguyen (2010). Risk and time preferences: Linking experimental and household survey data from vietnam. *American Economic Review 100*(1), 557–71.

Ton, G., W. Vellema, S. Desiere, S. Weituschat, and M. D'Haese (2018). Contract farming for improving smallholder incomes: What can we learn from effectiveness studies? *World Development 104*, 46–64.

Tversky, A. and D. Kahneman (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, Probability, and Human Decision Making*, pp. 141–162. Springer.

Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty 5*(4), 297–323.

Wang, H. H., Y. Wang, and M. S. Delgado (2014). The transition to modern agriculture: Contract farming in developing economies. *American Journal of Agricultural Economics 96*(5), 1257–1271.

A Tables

Table 7: Observations in contracting stages and treatments

	correct guess		underconfident		overconfident		Total	
Contracting								
total producing (seller)	17	(7.02%)	59	(24.38%)	166	(68.60%)	242	(100%)
offers (seller)	16	(6.93%)	56	(24.24%)	159	(68.83%)	231	(100%)
acceptances (buyer)	16	(8.60%)	45	(24.19%)	125	(67.20%)	186	(100%)
data used in analysis[1]	14	(8.14%)	42	(24.42%)	116	(67.44%)	172	(100%)
Treatments								
control	11	(64.71%)	42	(71.19%)	68	(41.72%)	121	(50.63%)
control (data used)[1]	10	(71.43%)	32	(76.19%)	49	(42.24%)	91	(52.91%)
anchor	6	(35.29%)	17	(28.81%)	95	(58.28%)	118	(49.37%)
anchor (data used)[1]	4	(28.57%)	10	(23.81%)	67	(57.76%)	81	(47.09%)

[1] Our data set that we use in the analysis comprises 172 observations in total (91 in the control and 81 in the treatment condition). Although we collected data of 186 sellers with accepted contracts, we had to drop 14 observations due to lost data from power outages in the field.

B Figures

Table 8: 1st stage results for Table 3

	Over-/underestimation	
	(2)	(4)
Anchor	2.681***	2.873***
	(0.662)	(0.727)
Shock	0.040	0.667
	(0.683)	(0.709)
Performance	-0.373***	-0.432***
	(0.087)	(0.093)
Constant	5.732	4.620
	(1.354)	(5.307)
Observations	172	172
R^2	0.181	0.254
Covariates[a]	NO	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[a]Covariates: age, income, share offered in ultimatum game, loss aversion, risk-seeking, SVO angle, guilt aversion, dummy for female, dummy for experiment's location (Kumasi or Accra), income per month (in real life), age in years, dummy for enrollment in an econ program.

In Model (4) the covariates guilt aversion, and subjects' income (in real life) are significantly positively correlated with over-/underestimation ($p < 0.10$)

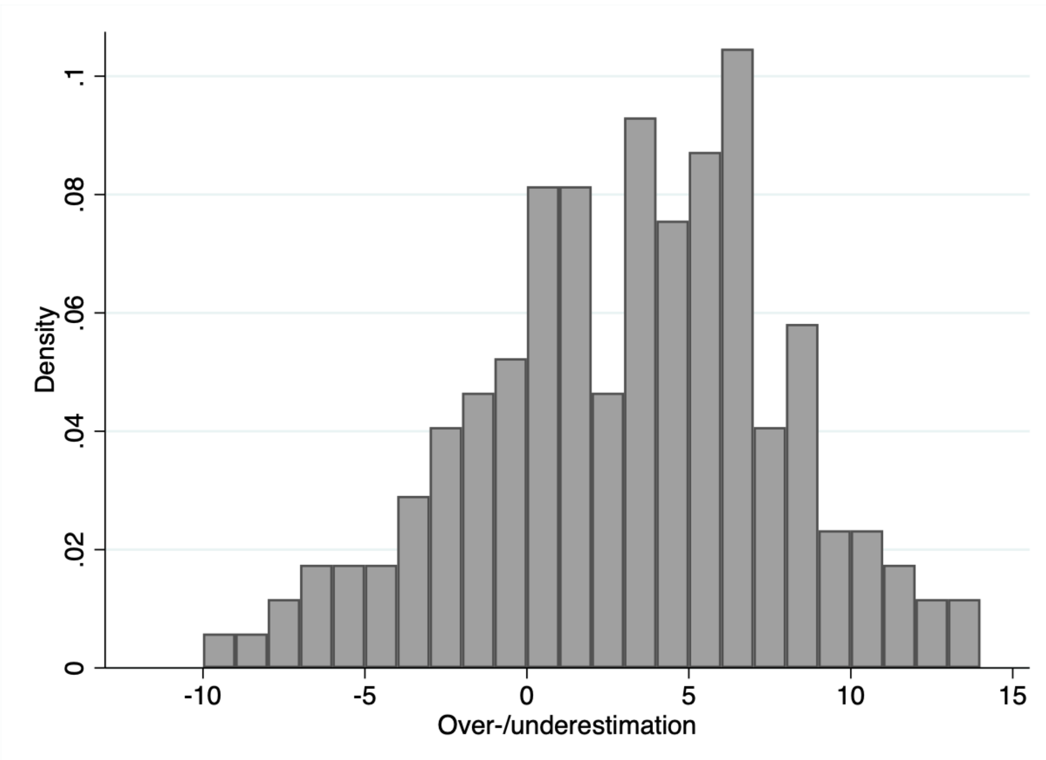

Figure 3: Histogram of over-/underestimation (N=172)