JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

GGS Giessen Graduate Centre for Social Sciences, Business, Economics and Law

BSFA Working Paper Series

The Effect of Compensation Caps on Risk-Taking

Niklas Kreilkamp
Sascha Matanovic
Friedrich Sommer
Arnt Wöhrmann

Working Paper No. 7 / August 2018

BSFA | Behavioral and Social Finance & Accounting
A research section at the Giessen Graduate Centre for Social Sciences, Business, Economics and Law

GGS | Justus Liebig University
Bismarckstrasse 22 | 35390 Giessen

Tel. +49 641 99 21370 | Fax +49 641 99 21379
bsfa@ggs.uni-giessen.de | www.uni-giessen.de/ggs/bsfa

The Effect of Compensation Caps on Risk-Taking

Niklas Kreilkamp
Giessen University (Germany)
Niklas.Kreilkamp@wirtschaft.uni-giessen.de
Licher Str. 62,
35394 Giessen, Germany

Sascha Matanovic
Giessen University (Germany)
Sascha.Matanovic@wirtschaft.uni-giessen.de
Licher Str. 62,
35394 Giessen, Germany

Friedrich Sommer
University of Bayreuth (Germany)
Friedrich.Sommer@uni-bayreuth.de
Nürnberger Str. 38
95448 Bayreuth, Germany

Arnt Wöhrmann
Giessen University (Germany)
Arnt.Woehrmann@wirtschaft.uni-giessen.de
Licher Str. 62,
35394 Giessen, Germany
(corresponding author)

Abstract

Compensation systems that restrict the earnings potential of managers (caps) are widespread across the business world and have gained even more importance since the financial crisis. We employ an experiment to examine the effect of compensation caps on risk-taking. Rational choice theory predicts that caps should only restrict risk-seeking managers from taking undesired risk and should not affect the decisions of risk-averse managers. We predict and find that risk-averse managers, who—according to their risk preferences—should not be affected by the cap, also decrease their risk-taking when their compensation is capped. This effect is magnified when justification pressure is high. Before considering differences in ex-ante risk preferences, we replicate prior research and show a general risk-decreasing effect of caps. Our results have important implications for theory and practice because we show that it is important to differentiate between risk-averse and risk-seeking individuals when anticipating the consequences of compensation caps. Firms should be aware that the implementation of a capped compensation system might lead to adverse consequences, as managers—particularly risk-averse ones—might take too little risk to exploit business opportunities with a balanced opportunity/risk relationship.

Keywords: Risk-taking, risk behavior, justification, compensation caps

JEL: M12; M41; M52

We thank Joyce Tian (discussant), Ta-Tung Cheng (discussant), the attendees at the 2018 American Accounting Association Annual Meeting and the attendees at the 2018 Canadian Academic Accounting Association Annual Conference as well as workshop participants at the University of Bayreuth for their helpful comments and suggestions. We are grateful for financial support by the Chartered Institute of Management Accountants.

I. Introduction

Managers must trade off the risk and rewards of decision alternatives in their daily business. Usually, alternatives that offer higher payoffs are associated with higher levels of risk. Compensation caps are regarded as a potential mechanism to limit the risk-taking of managers and other employees whose compensation depends on the realized payoff from a certain decision alternative (Hartmann and Slapničar 2015). Risk-taking is assumed to be reduced by a cap because higher and riskier payoffs do not translate into compensation beyond a level defined by the cap.

In response to excessive risk-taking observed during the financial crisis, lawmakers from the U.S. and Europe focused on compensation caps as an effective means to manage risk-taking (Murphy 2013; Asai 2016). However, caps should only affect high-risk choices—usually made by risk-seeking individuals—but not low-risk choices (usually made by risk-averse individuals) as business decisions always require a certain level of risk-taking.

In this paper, we examine whether compensation caps work as predicted. More precisely, we investigate whether compensation caps reduce risk-taking and whether the risk-reduction effect also materializes, though unintendedly, for managers who prefer low risk. In addition, we examine whether compensation caps and the requirement to justify one's decision, a popular alternative for managing risk-taking in firms, interact.

Prior literature provides ample evidence that compensation caps are widespread in the business world, even before the financial crisis (Murphy 2001; Jansen, Merchant, and van der Stede 2009). For instance, Murphy (2001) finds that of 177 large U.S. corporations, more than 80% employ caps for executive bonuses. Jansen et al. (2009) provide evidence that the use of caps is not limited to top executive compensation. The expected risk-reduction effect of caps, however, has motivated policy makers to require caps by law and not leave this decision to firms. For instance, U.S. firms receiving government funding to prevent bankruptcy during the financial crisis were required to implement caps in their compensation systems under the Troubled Asset Relief Program (TARP) (Garner and Kim 2010). Similarly, the European Commission and the European Parliament passed a law, i.e., the Capital Requirement Directive IV (CRD IV), to limit performance-contingent compensation to a predefined percentage of managers' fixed compensation (European Parliament and the Council 2013).

According to economic theory, i.e., rational choice theory, caps should limit the risk-taking of individuals with a preference for high risk. A cap—if present—virtually restricts the choices of these individuals. However, caps should not affect employees who prefer low levels of risk because their choices are in fact not restricted by the cap (Neumann and Morgenstern 2007). Rational choice theory argues that a preference for or against a decision alternative does not depend on the presence or absence of other (irrelevant) alternatives. Therefore, individuals who prefer a risk level below the level defined by the cap should make the same risk decision irrespective of whether a cap is present. This is because the cap does not reduce the *relevant* alternatives for these individuals. This principle, which is usually termed independence of irrelevant alternatives, implies that the decision maker has a complete preference order of all options. Hence, given a defined set of choices, the decision maker always selects the alternative that is ranked highest in that order. We use psychological theory to predict that even though caps—in general—work as predicted and reduce risk-taking, risk-averse individuals also reduce their risk-taking if a cap is present. Thus, caps are dysfunctional for some employees.

Risk is ubiquitous in decision making, and taking risks is essential for firms' success (Charness, Gneezy, and Imas 2013). However, avoiding high risk at the cost of taking too little risk is undesirable. This is particularly true as research shows that a preference for low levels of risk (risk aversion) is common among managers and—assuming that a firm strives for risk neutrality—

needs to be counteracted to adequately balance risk and return (Jensen and Meckling 1976; Low 2009).

Bebchuk and Spamann (2010) point out that compensation caps alone are not a solution to high or excessive risk-taking, as these authors identify unintended consequences of caps, such as lower effort and budget-gaming. Consequently, we do not investigate compensation caps in isolation, but we also examine whether caps and justification pressure (low vs. high) interact. Since the financial crisis, accountability for decision making in within organizations has grown (Tetlock, Vieider, Patil, and Grant 2013). Firms might use justification as an alternative measure to counter undesired levels of risk-taking (Lerner and Tetlock 1999; Lefebvre and Vieider 2013). Hence, we investigate the effectiveness of justification depending on the presence of cap.

Research on compensation caps is still scarce. Dittmann, Maug, and Zhang (2011) is one of the few studies in this field. These authors investigate different kinds of CEO pay restrictions and report several unintended consequences related to restrictions on ex post realized pay (e.g., stocks or options). These include, e.g., that CEOs earn on average more, are rewarded more for only average performance, and become more risk-averse. Asai (2016) examines the impact of caps on banks' risk-taking. He finds that caps reduce the risk-taking of executives but also increase underinvestment; that is, bank managers forego risky but profitable investment opportunities. Further, Jokivuolle, Keppo, and Yuan (2015) examine the effects of caps and extended bonus accrual periods. They show that extending bonus accrual periods alone does not decrease risk-taking, while a tight bonus cap does. Further, they find that a cap equal to the fixed salary of an employee decreases risk-taking by 13%. Kleymenova and Tuna (2016) examine the consequences of regulating executive compensation at financial institutions by investigating the introduction of the UK remuneration Code SYSC19C and the EU's CRD IV. They find that regulated investment services firms (so-called BIPRU firms) become less risky and show lower default risk under the

UK remuneration code. This finding lends support to the intended purpose of the UK regulation. While the studies discussed so far use archival data, Hartmann and Slapničar (2015) are the first to provide experimental evidence. These authors investigate the impact of negative, capped and deferred bonuses on risk-taking. They find that bonus caps reduce risk-taking more compared to a linear compensation system without upper or lower bounds. However, the economic and/or behavioral mechanism behind the risk-reducing effect of capped compensation systems remains unclear.

In this study, we predict that compensation caps not only reduce the risk-taking of risk-seeking managers (as intended) but also even further lowers the risk-taking of risk-averse managers, who in fact should not be affected by the cap, according to rational choice theory. We argue that caps are interpreted as a signal that the firm prefers a level of risk below a predefined threshold. Therefore, employees perceive the level of risk associated with the cap as “extreme”. As individuals try to avoid extreme options (extremeness aversion) and consider the whole range of alternatives (which is compressed when a cap is present) in their evaluation, risk-averse individuals also choose lower levels of risk. As mentioned above, this behavior is not in line with economic theory, as the cap does in fact not restrict the “natural” choice of these decision makers. For example, even when there is no cap, managers who prefer low levels of risk ignore high-risk options and decide on a low-risk (and low-return) alternative. When there is a cap, however, and the high-risk alternatives are thus removed, managers alter their decision and reduce the risk further. This is because a low-risk decision appears riskier under a capped than an uncapped compensation contract because the level set by the cap signals the most extreme option in terms of risk.

We predict that the risk-reduction effect induced by a cap is even higher if managers who prefer low levels of risk perceive high compared to low pressure to justify their decision. Today, justifying a decision to others is part of many organizational processes (Libby, Salterio, and Webb 2004; Pollmann, Potters, and Trautmann 2014), and justification pressure has been shown to affect the risk-taking decisions of managers (Schedlinsky, Sommer, and Wöhrmann 2018). The higher the pressure to justify a decision, the more likely it is that individuals choose an alternative that is easier to justify. Psychological research on the compromise effect argues that decision makers who must decide among several alternatives differing in at least two attributes (e.g., risk and return) choose a "middle option" (Simonson 1989). The idea of choosing the middle option is based on the assumption that decision makers have several motives for justifying their choices. The middle alternative is perceived as a reasonable compromise between the advantages and disadvantages of the other alternatives and is therefore the easiest to justify to oneself or others (e.g., superiors). This is in line with extremeness aversion, since alternatives in the bottom and upper parts of the choice set are rarely selected. In this regard, we predict that individuals who prefer low-risk options have a stronger risk focus (as opposed to a return focus) and hence interpret justification as an even stronger indication to reduce risk under the capped contract.

To test our predictions, we conduct an experiment with 447 participants using Amazon Mechanical Turk (MTurk). The procedures adhere to the guidelines for using MTurk in behavioral accounting research suggested by Buchheit, Doxey, Pollard, and Stinson (2018). We manipulate the *presence of a compensation cap* at two levels (*no-cap*/*cap)* and the decision maker's *justification pressure* at two levels *(low pressure/high pressure)*. The main task used to measure risk-taking is the bomb risk elicitation task (BRET task) suggested by Crosetto and Filippin (2013).

As predicted, we find that compensation caps lead to lower risk-taking. Most importantly, however, we find that in contrast to economic theory, risk-averse individuals (i.e., individuals choosing alternatives usually not restricted by the cap) also reduce their risk-taking when a cap is present. Further, the cap-induced decrease in risk-taking is more pronounced when justification pressure is high than when it is low.

This study contributes to accounting theory and practice. With respect to practice, our findings have important implications for the design of compensation systems. First, firms should be aware that compensation caps might lead to adverse consequences; i.e., risk-taking is reduced below reasonable levels. Second, we inform firms that build on accountability as a solution to the issue of undesired levels of risk-taking (Lerner and Tetlock 1999; Lefebvre and Vieider 2013). More precisely, we find that high justification pressure in the absence of a cap motivates risk-averse managers to take more risk than if justification pressure is low (Pollmann et al. 2014; Pahlke, Strasser, and Vieider 2012; Weigold and Schlenker 1991). However, under a cap, high justification pressure leads to less meaningful risk-taking and thus negatively affects firm performance.

From a theory perspective, we add to the scant research on the effect of caps on risk-taking. While it is well established that compensation caps reduce effort, the risk-taking effect is less clear. Prior experimental research examines the *overall* effect of capped compensation systems (Hartmann and Slapničar 2015). We complement this stream of research by investigating the effect of caps dependent on individuals' risk preferences and showing dysfunctional effects for risk-averse individuals. More precisely, we show that the overall risk-reducing effect of caps is to some extent due to the unintended behavior of risk-averse individuals. Hence, we also add to prior research by Sprinkle, Williamson, and Upton (2008), who investigate risk-taking under different compensation schemes, but focus on floors instead of caps. Further, for our theory development, we build on the compromise effect that is rooted in marketing research, i.e., consumer behavior, and demonstrate its implications for accounting. In detail, we show that extremeness aversion, and consequently the compromise effect, are important to consider when designing compensation systems.

In addition to our contribution to understanding the effects of caps on risk-taking, we investigate how different types of justification—an alternative management control practice to manage risk-taking—interact with compensation caps. While research on risk and accountability recommends the use of justification to improve risk-taking behavior (Lerner and Tetlock 1999; Lefebvre and Vieider 2013; Vieider 2009), we clarify that this mechanism can lead to undesired consequences under a capped compensation system.

The remainder of this paper proceeds as follows. Section II discusses the hypothesis development. Section III describes the experimental method, while Section IV presents the results. Section V concludes the paper.

II. Development of Hypotheses

The Overall Effect of Compensation Caps on Risk-taking (H1)

Murphy (2013) suggests compensation caps as an effective means to limit (excessive) risk-taking in firms. While little empirical research on the effectiveness of caps exists, this research generally confirms that caps reduce risk-taking. However, while Dittmann et al. (2011) confirm this intended result, they also report numerous unintended consequences when restricting realized pay in the context of stocks and options. These authors report that CEOs earn on average more, are rewarded more for average performance, and become more reluctant to accept risks if pay restrictions are present. Asai (2016) adds that even though compensation caps reduce the risk-taking of managers, such caps also lead to severe underinvestment; i.e., managers forego risky but profitable investment opportunities. In an experimental setting, Hartmann and Slapničar (2015) investigate a linear compensation system with a loss and a gain domain and a capped compensation system. The linear compensation system has a cap at a defined point in the gain domain and a floor at zero so that the participants do not incur losses. These authors rely on prospect theory to predict that such linear compensation systems (mixed domains) lead to lower risk-taking than a capped compensation system (only positive domain). However, their results show the opposite, and the hypothesis that non-capped compensation systems lead to lower risk-taking than capped compensation systems is rejected. Hence, this experimental research finds that compensation caps reduce risk-taking. The results of prior research are in line with rational expectations. When managers reach a compensation cap (or come close to it), taking more risk is irrational as there is no reward for more risk. This leads to our first hypothesis.

H1: Capped compensation systems decrease risk-taking.

The Effect of Compensation Caps on Risk-taking for Low Risk Takers (H2)

While H1 focuses on the general effect of compensation caps on risk-taking, H2a and H2b predict how compensation caps affect the risk-taking of managers who generally prefer low levels of risk. From a rational perspective, these decision makers should not alter their level of risk if a cap is present. This prediction rests upon the assumption that decision alternatives with high potential payoffs also entail high risk and that the ex-ante risk preference lets risk-averse decision makers choose alternatives with potential payoffs below the cap.

Technically speaking, adding a cap to a manager's compensation contract decreases the number of relevant options among which the manager can choose. This is because the upside potential of high-risk options is limited when a cap is implemented. However, according to economic theory, managers who prefer low levels of risk anyway (leading to potential outcomes below the cap) should not be affected when a cap is introduced that makes high-risk options (even more) unattractive. More precisely, rational choice theory states that a preference between several options does not depend on the availability of other irrelevant alternatives. This principle is also referred to as the independence of irrelevant alternatives (IIA) (Luce 2005; Tversky and Simonson

1993). Further, the regularity axiom of probability theory states that the probability of choosing one alternative over another should not be affected by the addition or removal of other alternatives (Luce 1977). Taken together, caps should not affect the risk-taking behavior of managers who prefer low levels of risk if the cap is sufficiently high, that is, if the cap is set at a level where high-risk alternatives (but not low-risk alternatives) become (even more) unattractive.

However, we rely on psychological theory that predicts that caps (unintendedly) affect risk-averse managers. This follows from research on the compromise effect mainly discussed in marketing. Broadly speaking, the compromise effect predicts that decision makers who must decide among several alternatives that differ in at least two attributes (e.g., quality and price) choose a middle option that appears to be a good comprise between all relevant attributes (Simonson 1989). The attributes, however, must be symmetric for the compromise effect to materialize (Tversky and Simonson 1992). For instance, choosing an alternative with a higher quality inevitably leads to a higher price. Accordingly, trying to increase the advantage of one attribute (e.g., quality) by choosing an alternative of higher value also magnifies the disadvantage of the other attribute (e.g., price). The idea of choosing the middle option is based on the assumption that decision makers have several motives for justifying their choices. The middle alternative appears to be a reasonable "compromise" or "a good trade-off" between the advantages and disadvantages of the other alternatives and is therefore easy to justify to oneself and others.

Tversky and Simonson (1992) account for so-called "extremeness aversion", whereby people seek a compromise between disadvantages and advantages and thus choose the middle option. This extremeness aversion hypothesis builds on the presence of loss aversion: outcomes below a reference point (losses) loom larger than outcomes of an identical magnitude above this reference point (gains) (Tversky and Kahneman 1991). Tversky and Simonson (1992) extend this notion to advantages and disadvantages defined in relation to other available alternatives independent of the reference point. Hence, disadvantages (Kahneman and Tversky 1981) have a stronger effect than the respective advantages, and people prefer the middle option with relatively small disadvantages compared to other available options. In particular, loss aversion, and therefore extremeness aversion, is stronger the stronger the risk aversion is (Kahneman and Tversky 1979; Chuang, Cheng, Chang, and Chiang 2013).

The compromise effect violates the IIA principle. According to this principle, the preference order of a decision maker who has to decide between two options, A (e.g., low risk, low return) and B (e.g., medium risk, medium return), should not be affected when a third option, C (e.g., high risk, high return), is added (or removed) that is outside her preferences. In other words, if the decision maker prefers A to B or vice versa, her preference should not change when alternative C is added (or removed).[1]

The compromise effect has important implications for behavior under capped compensation contracts. Managers who prefer low levels of risk choose a low-risk (and low-return) alternative and ignore high-risk options even if there is no cap. However, when compensation is capped and thus high-risk alternatives are virtually removed, managers alter their decisions. This is because a low-risk decision in the absence of a cap appears riskier in the presence of a cap since all high-risk options are removed, and the most extreme option now available is in fact a medium-risk option. This is formally stated by H2a.

[1] In marketing research, the alternatives usually investigated contain non-numeric (e.g., quality) and numeric attributes (e.g., price). Tversky and Simonson (1992) investigate choice behavior with alternatives having the attributes of price and quality in order to make it more difficult for the decision maker to determine a preference order based on attribute weights and values alone (i.e., one numeric attribute and one non-numeric attribute instead of two numeric attributes). In our setting, individuals can now build preferences based on values and hence make calculations (e.g., use risk and return to calculate the expected value) (Simonson 2014; Neumann, Böckenholt, and Sinha 2016). As two numeric attributes are easily measurable, this allows an unambiguous interpretation of the trade-off.

H2a: Managers who prefer low levels of risk reduce their risk-taking when their compensation is capped.

When making a decision, individuals strive to be able to justify their decision to themselves (internal justification – low pressure) to enhance their self-esteem. Being unable to justify a decision results in negative consequences, such as a negative emotional state stemming from cognitive dissonance (Festinger 1957; Holland, Meertens, and van Vugt 2002; Lerner and Tetlock 1999) or regret (Kahneman and Tversky 1981; Samuelson and Zeckhauser 1988; Simonson 1989).

If a decision is justified to others, such as superiors (external justification – high pressure), two forms of external justification may materialize. The first form occurs when justification is explicitly required, e.g., by a superior (Libby et al. 2004). The second form occurs if a third party observes the decision, and the decision maker wants to be perceived as competent (Tetlock 1985).

The major difference between internal (low justification pressure) and external justification (high justification pressure) is the fact that individuals are more familiar with their own preferences than with those of others. This is important for firms where individuals' peers and superiors, as well as the firm itself (e.g., via a code of conduct), do not have identical risk preferences. Thus, choosing a middle option when external justification is required provides universal reasons and is the easiest to justify. This is because a middle option appears to be an acceptable compromise between all the advantages and disadvantages of the alternatives. Hence, justification in any form is a vital component of the compromise effect.

Managers who have to justify their decision to others are therefore even more likely to choose the middle option as this protects them against potential criticism. Consequently, when individuals expect that they will be required to provide external justification, the compromise effect is stronger for external justification (high pressure) than for internal justification (low pressure) (Chernev 2005; Neumann, Böckenholt, and Sinha 2016; Simonson 1989). In particular, in cases where compensation is not capped, risk-averse managers take more risk when external justification is required than when internal justification is required based on the assumption that the firm tries to balance risk and return and hence strives for medium levels of risk (Weigold and Schlenker 1991). However, in cases of capped compensation, managers take less risk when external justification is required than when internal justification is required.[2] This leads to H2b (see Figure 1).

H2b: Managers who prefer low levels of risk reduce their risk-taking under a cap even more if justification pressure is high than if justification pressure is low.

[Place Figure 1 about here]

III. Experimental Method

Experimental Design and Manipulations

To test our hypotheses, we conduct an online experiment using MTurk. The experiment was programmed and conducted using the software package SoPHIE (Hendriks 2012). We manipulate the *presence of a compensation cap* at two levels (*no-cap*/*cap)* and *justification*

2 We do not predict any main effect of accountability, particularly justification pressure in the context of cap analysis, for two reasons. First, justification remains a major indispensable mechanism for the compromise effect from a theoretical perspective. Hence, we investigate the level of justification pressure that is necessary to evoke this effect. Second, the risk-taking of risk-averse managers increases with the level of justification in cases where no capped incentive systems are established. In contrast, the risk-taking of risk-averse managers decreases with the level of justification in cases where a capped incentive system is established. As a result, a main effect for justification cannot be observed.

pressure at two levels *(low pressure/high pressure)*. This results in a 2×2 full factorial between-subjects design.

The bomb risk elicitation task (BRET task) (Crosetto and Filippin 2013) is our main task and is described in more detail below. In this task, participants decide to collect between 0 and 100 virtual boxes knowing that a bomb is hidden in one of the 100 boxes. Each box collected increases their compensation by $0.03. If, however, they collect the bomb, the compensation from this task is zero. Thus, the compensation (risk) increases by 0.03 USD (1%) per box. Importantly, participants do not learn immediately whether they have collected a box with the bomb. Rather, they need to finish the task before and later receive the summary on their compensation that informs them about whether they collected the bomb.

In the *no-cap* treatment, participants have to choose to collect between 0 and 100 boxes and receive $0.03 per box if they do not collect the bomb. In the *cap* treatment, participants also choose to collect between 0 and 100 boxes and receive $0.03 per box if they do not collect the bomb. However, the variable compensation in the cap treatment is limited to $1.5, which equals the compensation for collecting exactly 50 boxes.[3]

Justification pressure (low pressure/high pressure) is manipulated as follows: In the *high-pressure* treatment, participants have to justify their decision regarding how many boxes they collect. More precisely, participants have to enter a written justification in a textbox and are informed that they might have to answer a follow-up question concerning this justification.[4] In the

[3] Participants are paid for taking additional risk only until the alternative with the highest level of expected value is reached. With regard to the level of the compensation limit (cap), various alternatives appear possible. We choose to set the compensation limit at the alternative with the highest expected value (50 boxes) because firms are also likely select the optimal level of risk as an upper boundary of compensation. This approach allows a relatively high level of external validity.

[4] Eleven participants from the justification treatments (5% of all relevant participants) are randomly selected and receive a follow-up question ("We would kindly ask you to clarify the justification you provided during our experiment. Please briefly answer the following question: Were you more motivated by the desire to avoid risk (getting the bomb) or by the desire to earn high compensation?").

low-pressure treatment, no justification is required. However, previous research argues that individuals justify decisions to themselves (Simonson 1989). This is why we consider this treatment a low-pressure treatment. Participants are randomly assigned to one of the four treatment conditions.

To measure ex-ante risk preferences, we include a risk-elicitation instrument used in prior research (e.g., Sprinkle et al. 2008). For each of 15 scenarios, participants have to choose between a safe payment of \$0.75 and a lottery that pays either \$1.50 with a probability of π or \$0 with a probability of ($1-\pi$). The probability π decreases from 85% (state of nature 1) to 15% (state of nature 15) in 5% increments. The later participants switch from the lottery to the safe payment, the more risk-seeking they are. We refer to participants who switch from the lottery to the safe payment before scenario 8 as risk-averse participants and at scenario 8 as risk-neutral participants. We refer to the other participants as risk-seeking participants.[5]

Task

As stated above, the main task is the BRET task by Crosetto and Filippin (2013). At the beginning of the task, the participants see a screen with 10×10 cells, each cell representing a box. Two buttons, "Start" and "Stop", are located below the boxes. The task begins when a participant clicks the "Start" button. Each second, one cell is automatically deleted from the screen, which means that one box is collected per second. The deletion process starts from the top-left corner (Figure 2). The number of boxes collected and the corresponding—however provisional—compensation are continuously displayed. Each time a box is collected, the participant's

[5] Measuring ex-ante risk preferences and risk-taking behavior with two different tasks has several advantages over using the same task twice in a within-subject design. Subjects are less likely to connect the two tasks to infer that the main difference between the ex-ante measurement and the main task is the introduction of the cap. Using the same task twice would result in a within-subject design that would clearly investigate the effects of introducing a cap (instead of the existence of a cap). Although this might be a fruitful research question for the future, this design does not match our theory and is therefore not implemented.

provisional account is credited with additional \$0.03. Such earnings are only potential, however, because one of these boxes hides a bomb that destroys everything that has been collected.

[Place Figure 2 here]

The participants know that the bomb can be in any box (boxes 1-100) with equal probability and that they are informed that the earnings displayed during the task are only provisional. However, they do not know the position of the bomb. At the end of the experiment, a random mechanism determines whether the bomb has been collected. This feature of the task is important because we would otherwise not be able to avoid the truncation of the data that would occur in the case of a real-time notification. The participant can stop the drawing process and thereby finish the task at any time by clicking on the "Stop" button.

Notably, a risk-neutral decision maker would collect 50 boxes to obtain the highest expected value.[6] From a theoretical perspective, participants can choose between 101 lotteries fully described in terms of outcomes and probabilities by a single factor, i.e., the number of boxes collected. The expected value of these lotteries is equal to $\alpha \times (\beta - 0.01 \times \beta^2)$, a bow-shaped function with a maximum at $\beta = 50$ and equal to zero for $\beta = 0$ and $\beta = 100$. Here, α is equal to the dollar amount paid per box, and β represents the number of boxes collected.

Procedures

Figure 3 depicts the experimental procedure. When the experiment started, the instructions were displayed. These instructions contained information about the compensation and the BRET task. After that, the participants completed the ex-ante risk-elicitation task described earlier. Afterwards, they received detailed instructions about the main task of the experiment and were

[6] Accordingly, a risk-averse decision maker would select fewer than 50 boxes, and a risk-seeking decision maker would select more than 50 boxes.

required to take a quiz to demonstrate an adequate understanding of the experimental procedure and the computation of their compensation. If they answered all questions correctly, they could proceed with the experiment and start the BRET task. At the end of the experiment, the participants completed a set of questions on a post-experimental questionnaire and were informed about their compensation. The compensation consisted of a fixed participation fee of $1.00, a payoff from the ex-ante risk-elicitation task (either $1.50 or $0, depending on their choice and the realized state of nature), and their earnings from the BRET task detailed above. On average, the compensation was $2.41, and the experiment lasted approximately 15-20 minutes. The compensation in this study was above the average MTurk reservation wage of $1.38 per hour (Horton and Chilton 2010).

[Place Figure 3 here]

Participants

A total of 447 individuals from the MTurk internet marketplace were recruited for the experiment through a publicly advertised Human Intelligence Task (HIT). MTurk provides an easily accessible and cost-effective participant pool that has been proven to provide reliable data (Brasel, Doxey, Grenier, and Reffett 2016; Paolacci, Chandler, and Ipeirotis 2010). Furthermore, MTurk workers are more diverse and representative of the U.S. population than undergraduate students (Buhrmester, Kwang, and Gosling 2011; Farrell, Grenier, and Leiby 2017; Paolacci et al. 2010). They also exhibit a comparable level of intelligence as traditional research participants (Buchheit, Dalton, Pollard, and Stinson 2016). Prior studies in accounting have successfully used MTurk participants (Koonce, Miller, and Winchel 2015; Christ and Vance 2017; Grenier, Pomeroy, and Stern 2015; Rennekamp 2012; Asay, Elliott, and Rennekamp 2017). Recent research by Farrell et al. (2017) replicates three existing accounting studies using participants from MTurk with similar results, further strengthening the reliability of the population.

As suggested by Buchheit, Doxey, Pollard, and Stinson (2018), we screened the population of interest, included four attention checks, and provided disclosure at the very beginning of the experiment to ensure subject quality. Workers were eligible to participate in the study as long as they had a historical HIT approval rating of at least 95%, proved that they were 18 years or older and were based in the U.S. (Brown, Sprinkle, and Way 2017; Christ and Vance 2017). We included several questions to ensure that MTurk workers understood the experiment and were attentive when answering the post-experimental questionnaire. Eight out of 447 participants failed one or more attention check questions (Meade and Craig 2012). Excluding these participants led to inferentially inferentially identical results (see footnotes 8 and 14).

The participants' average age was 36.5 years, 223 participants (49.89%) were male, and 70.02 percent had attended college. Using the ex-ante risk-elicitation task, we found that 73.15% of the participants are risk-averse, 9.62% risk-neutral and 17.23% risk-seeking. There were no significant differences across conditions for age (p = 0.34, two-tailed, Kruskal-Wallis test), sex *(p* = 0.17, two-tailed, chi-square test), education *(p* = 0.25, two-tailed, chi-square test), statistical knowledge (p = 0.93, two-tailed, Kruskal-Wallis test) or risk preferences[7] (p = 0.69, two-tailed, Kruskal-Wallis test). Hence, we concluded that the randomization was successful.

IV. Results

Descriptive Statistics

Table 1 depicts the mean and the standard deviation for our main dependent variable, *number of boxes collected*. As explained above, the number of boxes collected is our operationalization of risk-taking. The descriptive results for the number of boxes collected are in line with our hypotheses. Across all subjects, risk-taking decreases from 34.56 in the no-cap

[7] For the purposes of this randomization check, we use the scale of 1 to 15 from the lottery task.

condition to 30.73 in the cap condition. Risk-averse individuals collect 28.44 boxes in the no-cap condition but only 26.19 boxes in the cap condition. The decrease in risk-taking caused by the presence of a cap is stronger if justification pressure is high (from 28.90 to 24.14) than when it is low (from 27.99 to 27.92).

As discussed above, we employ a risk-elicitation instrument (before employing the main BRET task) to measure ex-ante risk preferences and classify participants as risk-averse, risk-neutral, or risk-seeking. Of those classified as risk-averse, 52 collected more than 49 boxes in the no-cap treatment (untabulated). Thus, these subjects exhibited risk-seeking behavior in the BRET task. Two explanations for this incongruent behavior appear equally likely. First, even in the absence of any manipulation, the fact that individuals are not perfectly identically classified by two different risk-elicitation methods is not new (Lönnqvist, Verkasalo, Walkowitz, and Wichardt 2015). Second, individual risk preference does not have to translate into actual risk-taking behavior because it depends on the situation. In our case, the ex-ante risk preferences cannot be affected by the manipulation, whereas the measurement in the BRET task is necessarily affected. To rule out that these subjects impact our results and to provide a conservative test for our theory, we eliminate these 52 observations for the tests of H2a and H2b. Therefore, we rule that the risk-reduction effect observed for risk-averse subjects under the cap contract is in fact the result of (risk-averse) subjects in the cap treatment taking too high levels of risk, i.e., collecting more than 50 boxes.

[Insert Table 1 about here]

Hypothesis Tests

H1 predicts that caps reduce risk-taking. The descriptive statistics in Table 1 reveal that a cap reduces the number of boxes collected from 34.56 to 30.73 across all subjects, that is, irrespective of risk preference. To formally test the predicted direction of H1, we use analysis of variance (ANOVA) with the number of boxes collected as the dependent variable. Table 2 contains the ANOVA results that support H1 ($F = 8.37$, $p < 0.01$, one-tailed). Therefore, we conclude that a capped compensation system decreases risk-taking.[8]

[Insert Table 2 about here]

H2a predicts that caps reduce risk-taking even among risk-averse managers, though from a rational perspective, caps should not affect them. The descriptive statistics in Table 1 indicate that risk-averse subjects in fact collect fewer boxes in the cap condition (26.19) than in the no-cap condition (28.44). H2a is formally tested using ANOVA with the number of boxes collected as the dependent variable. The results are presented in Table 3, Panel A. As predicted by H2a, the effect of a cap on the number of boxes collected is significant ($F = 4.13$, $p = 0.02$, one-tailed). Hence, we conclude that the compensation cap further reduces risk-taking by managers with a level of risk that is already below the cap. Consequently, these managers are affected by the cap, although from a rational perspective (theory of rational choice) and oftentimes from the managerial perspective (balancing opportunity and risk), they should not be.[9]

To tie these results closer to the compromise effect, we further investigate the psychological mechanisms. The compromise effect results from extremeness aversion. In detail, risk-averse managers perceive the cap as an extreme value that they then try to avoid. That is, subjects in the cap treatment should perceive the value at the cap (i.e., 50 boxes) as riskier than subjects in the no-cap condition do. A Wilcoxon-Mann-Whitney test confirms that subjects in the cap condition

---

8 In testing H1, when excluding 8 out of 447 participants who failed at least one of our four attention checks, our results remain inferentially identical.

9 The cap is set at 50 boxes. A risk-neutral decision maker would collect 50 boxes to obtain the highest expected value. A risk-averse decision maker, however, would always select fewer than 50 boxes.

perceive the value at the cap as more extreme than subjects in the no-cap condition do (Z = -2.96, $p < 0.01$, two-tailed).[10]

H2b predicts that caps reduce the risk-taking of managers who prefer low levels of risk more if justification pressure is high than when it is low. The descriptive results in Table 1 and Figure 4 show that external justification slightly increases risk-taking in the no-cap condition (27.99 boxes for low justification vs. 28.90 boxes for high justification) and decreases risk-taking in the cap condition (27.92 boxes for low justification vs. 24.14 boxes for high justification). These results indicate that external justification strengthens extremeness aversion and, ultimately, the compromise effect.[11] The ANOVA results in Table 3, Panel A, confirm a significant interaction effect of cap and justification (F = 376.94, $p = 0.02$, one-tailed). To formally test the complete pattern predicted in H2b, we rely on contrast analysis (Buckless and Ravenscroft 1990). According to our prediction (depicted in Figure 1), we use contrast weights of +1 for no-cap/low justification, +3 for no-cap/high justification, -1 for cap/low justification, and -3 for cap/high justification. Table 3, Panel B, contains the test results, which support H2b (F = 7.09, $p < 0.01$, one-tailed).[12] To further validate the finding without relying on contrast analysis, we test whether the slopes for the effect of the presence of a cap differ depending on justification pressure (low/high) (Cohen 1983).[13] The effect size of capped compensation is higher ($z = 1.98$; $p = 0.01$, one-tailed) when justification is present (β_1 = -4.67; standard error (SE) = 1.71) than when it is absent (β_2 = -0.06; SE = 1.65).

10 The post-experiment questionnaire asked participants “How risky do you think is it to collect 50 boxes?” which was answered on a Likert scale from 1 (not risky at all) to 11 (very risky). Participants in the no-cap condition responded with 8.08 on average, while participants in the cap condition responded 8.51.

11 Subjects were asked “How risky do you think is it to collect 50 boxes?“ on the post-experimental questionnaire. We run an ANOVA identical to the analyses reported in Table 2 but with the response to this question as the dependent variable. We find a significant main effect ($p = 0.06$, one-tailed), indicating that the extreme values are considered riskier.

12 The results reported in Table 3, Panel B, for H2b are robust to the use of alternative contrast weights +2, +3, +1 and -6 (F = 8.97, $p < 0.01$, one-tailed) and +2, +3, +2, -7 (F = 8.96, $p < 0.01$, one-tailed).

13 The effect sizes and standard errors obtained are used to calculate the z-statistic as follows: $z = \frac{\hat{\beta}_1 - \hat{\beta}_2}{\sqrt{se(\hat{\beta}_1)^2 + se(\hat{\beta}_2)^2}}$

Therefore, we conclude that external justification strengthens the compromise effect and decreases the risk-taking of subjects who prefer low levels of risk.[14]

[Insert Table 3 about here]

[Insert Figure 4 about here]

Additional Analyses

In this section, we present further analyses to rule out alternative explanations for our results and to validate that our experimental results for H2a and H2b are driven by the proposed theoretical concepts.

With respect to alternative explanations, we first rule out that participants had a specific compensation goal in mind while working on the MTurk experiment, which could have represented well-established preferences and interfered with our results. Tversky and Simonson (1992, 292) conclude that "both tradeoff contrast and extremeness aversion are expected to have less impact in situations in which consumers have well-established preferences". Therefore, we asked our MTurk participants if they had a specific compensation target in mind during our main task.[15] If they did, this could have potentially interacted with our findings. In our case, when participants started the experiment with a target, they could have calculated the expected values for each box and hence chosen the box with the expected value that fit their target. Consequently, if they had a target (i.e., a well-established preference), extremeness aversion and, as a result, the compromise effect would

[14] The results stay inferentially identical when, in testing H2a and H2b, we exclude 3 out of 275 participants who failed at least one of our four attention checks. For the remaining non-risk-averse participants (risk-neutral and risk-seeking), we do not find a risk-reducing main effect for caps ($p = 0.18$, one-tailed), but we find a significant impact of justification pressure ($p = 0.02$, one-tailed) and a significant interaction effect ($p = 0.04$, one-tailed). In the absence of justification, capped compensation increases risk-taking on average from 31.44 to 33.64. In contrast, under justification pressure, capped compensation reduces risk-taking on average from 42.47 to 35.44.

[15] We asked the participants "Did you aim at earning a specific compensation from the bomb task?" They could only respond "Yes" or "No".

have had less of an impact. We therefore excluded participants who stated that they had such a compensation target in mind. Table 4 reports risk-taking only for risk-averse subjects (identical to the bottom part of Table 1, presented again for direct comparison) and risk-averse subjects without such a clear compensation target.

The descriptive statistics in Table 4 and Figure 5 show an even stronger decrease in risk-taking for subjects without a target when capped compensation is used. The number of boxes collected by risk-averse subjects decreases by 9 percent (from 28.44 to 26.19) if subjects with a compensation target are not excluded but by 20 percent (from 27.45 to 22.82) when they are excluded. Further, justification pressure leads to a sharper decrease in risk-taking when a cap is present than when it is absent. In detail, risk-taking decreases by 20 percent (from 28.90 to 24.14 collected boxes) under high justification pressure in the absence of a cap with all risk-averse subjects included, while risk-taking decreases by 38 percent (from 29.15 to 21.15 collected boxes) when we exclude participants with a compensation target in mind. This analysis indicates that the compromise effect is even more pronounced when subjects have no compensation target in mind (i.e., no well-established preferences). This is in line with the results and conclusion of Tversky and Simonson (1992).

[Insert Table 4 about here]

[Place Figure 5 about here]

Further, the results for H2a remain robust and become even more significant. Table 5 presents the corresponding hypothesis tests for risk-averse subjects without a target in mind. H2a states that a capped compensation system reduces the risk-taking of risk-averse managers, although they are not affected by the cap. Again, we use ANOVA to replicate the test results for H2a without subjects who stated that they had a compensation target in mind. Again, H2a is supported (F =

21.07, $p < 0.01$, one-tailed). Therefore, we conclude, in line with Tversky and Simonson (1992), that extremeness aversion and hence the compromise effect have less impact in situations with well-established preferences[16]

[Insert Table 5 about here]

Next, we use questions from the post-experiment questionnaire to further validate the compromise effect and the psychological mechanisms, i.e., extremeness aversion and justification, for subjects who prefer low levels of risk.

The compromise effect and the enhancement through justification work in both directions because both situations (risk increase in the no-cap condition and risk decrease in the cap condition) have (implicit) upper bounds (100 and 50 boxes, respectively). Hence, the participants should *not* differ in compromising between the two attributes of risk and return between the two conditions. Participants responded on a Likert scale from 1 (not at all) to 11 (very much) to the following question: "Did you try to compromise between risk and potential compensation?" Overall, participants responded with 8.68 on average, which is significantly different from the midpoint of 6 (T = 20.13, $p < 0.01$, two-tailed). However, we find no differences between the cap and no-cap conditions (Z = -1.45, $p = 0.15$, two-tailed). Hence, the compromise effect exists when compensation is capped as well as when compensation is not capped.

Moreover, to verify that external justification (high pressure) results in more pressure than internal justification (low pressure), the participants responded on a Likert scale from 1 (not at all) to 11 (very much) to the following question: "I felt as if I should thoroughly contemplate my

[16] Indeed, an ANOVA with risk-taking as the dependent variable along with capped compensation (present/absent) and the presence of a compensation goal (present/absent) shows an interaction of the two factors ($p = 0.02$, one-tailed). If participants have a compensation goal, capping the compensation has hardly any impact (mean risk-taking changes from 29.13 to 29.12). In the absence of a compensation goal, the cap reduces risk-taking from 27.45 to 22.82.

choices in the bomb task". Participants in the low-justification-pressure condition responded with 7.19 on average, while participants in the high-justification-pressure condition responded with 7.70 on average. The difference is statistically significant (Z = -1.83, p = 0.07, two-tailed). Consequently, we conclude that our operationalization of justification pressure was successful. In addition, Simonson (1989) and Tversky and Simonson (1992) argue that external justification strengthens extremeness aversion and therefore the compromise effect. Consequently, we asked participants on a Likert scale from 1 (not at all) to 11 (totally agree): "Did you try to balance the relative advantages and disadvantages connected with taking more or fewer boxes?" Participants in the low-justification-pressure condition responded 8.22 on average, while participants in the high-justification-pressure condition responded 8.49 on average. A Wilcoxon-Mann-Whitney test yields significant results (Z = -1.67, p = 0.09, two-tailed). Hence, we conclude that the urge to select the middle option increases when subjects have to justify their decision to others.

V. Conclusion

Compensation caps are widespread in the business world. Previous research documents that caps can be used to limit risk-taking among managers (Hartmann and Slapničar 2015; Jokivuolle et al. 2015; Asai 2016; Kleymenova and Tuna 2016). We investigate whether a capped compensation system influences the risk-taking behavior of only risk-neutral or risk-seeking managers or, contrary to the expectations of rational choice theory, it also influences the behavior of risk-averse managers. Additionally, we analyze whether the level of justification pressure matters.

We conduct an online experiment using MTurk, where participants make decisions under risk. We manipulate the presence of a compensation cap at two levels (no-cap/cap) and justification pressure at two levels (low pressure/high pressure).

We provide evidence that compensation caps lead to lower risk-taking. In alignment with prior studies, we also show that part of the reason for the reduction in risk-taking is that risk-averse participants take less risk when a cap is present (although not necessary given their ex-ante preferences). This reduction in risk-taking is magnified by high justification pressure.

Our findings have important implications for management accountants involved in the design of compensation systems. First, our results show that the implementation of a capped compensation system leads to unintended consequences. As companies strive to balance opportunity and risk, the reduction of risk-taking by risk-averse managers might counteract organizational objectives. Second, our results regarding justification imply that the combination of caps and high (external) justification pressure might aggravate the problem.

Future research might further explore this field of research. While our study demonstrates that differentiating between the ex-ante risk preferences of managers is essential when anticipating their risk-taking behavior under a capped compensation system, we did not focus on round effects or time effects. Future studies might also examine the effect of a compensation system change on risk-taking behavior in a within-subject design. Implementing a compensation cap after managers have previously worked under an unlimited compensation system might have different effects on their risk-taking behavior. Further, we did not investigate the effect of social status within a peer group on the relationship between compensation caps and risk-taking. It would be interesting to explore whether status concerns potentially lead to an increase in risk-taking, similar to prior findings showing that social status concerns have an effort-increasing effect on managers. In addition, a differentiation between risk and uncertainty could probably alter the results and hence be of interest for further investigation. Several studies have documented a relationship between risk-taking and effort. Finally, future research can also investigate the effect on effort, since fairness concerns, trust and reciprocity can also play a role when a cap is implemented by the government or by higher-level managers (Christ, Sedatole, and Towry 2012; Christ 2013).

References

Asai, K. 2016. Is Capping Executive Bonuses Useful? *International Monetary Fund* 16/196.

Asay, H. S., W. B. Elliott, and K. Rennekamp. 2017. Disclosure Readability and the Sensitivity of Investors' Valuation Judgments to Outside Information. *The Accounting Review* 92 (4): 1–25.

Bebchuk, L. A., and H. Spamann. 2010. Regulating Bankers' Pay. *Georgetown Law Journal* 98 (2): 247–287.

Brasel, K., M. M. Doxey, J. H. Grenier, and A. Reffett. 2016. Risk Disclosure Preceding Negative Outcomes. The Effects of Reporting Critical Audit Matters on Judgments of Auditor Liability. *The Accounting Review* 91 (5): 1345–1362.

Brown, J. L., G. B. Sprinkle, and D. Way. 2017. *The Effects of Multi-Level Group Identification on Intergroup Cooperation and Performance*. Working paper, Indiana University – Kelley School of Business.

Buchheit, S., D. Dalton, T. Pollard, and S. Stinson. 2017. *How Smart Are Online Workers? A Student versus MTurk Participant Comparison*. Working paper, The University of Alabama and Clemson University.

Buchheit, S., M. M. Doxey, T. Pollard, and S. R. Stinson. 2018. A Technical Guide to Using Amazon's Mechanical Turk in Behavioral Accounting Research. *Behavioral Research in Accounting* 30 (1): 111–122.

Buckless, F. A., and S. P. Ravenscroft. 1990. Contrast Coding: A Refinement of ANOVA in Behavioral Analysis. *The Accounting Review* 65 (4): 933–945.

Buhrmester, M., T. Kwang, and S. D. Gosling. 2011. Amazon's Mechanical Turk. A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science* 6 (1): 3–5.

Charness, G., U. Gneezy, and A. Imas. 2013. Experimental methods. Eliciting risk preferences. *Journal of Economic Behavior & Organization* 87: 43–51.

Chernev, A. 2005. Context Effects without a Context. Attribute Balance as a Reason for Choice. *Journal of Consumer Research* 32 (2): 213–223.

Christ, M. H., K. L. Sedatole, and K. L. Towry. 2012. Sticks and Carrots: The Effect of Contract Frame on Effort in Incomplete Contracts. *The Accounting Review* 87 (6): 1913–1938.

Christ, M. H. 2013. An Experimental Investigation of the Interactions among Intentions, Reciprocity, and Control. *Journal of Management Accounting Research* 25 (1): 169–197.

Christ, M. H., and T. W. Vance. 2017. Cascading controls. The effects of managers' incentives on subordinate effort to help or harm. *Accounting, Organizations and Society* 65: 20–32.

Chuang, S.-C., Y.-H. Cheng, C.-J. Chang, and Y.-T. Chiang. 2013. The impact of self-confidence on the compromise effect. *International Journal of Psychology* 48 (4): 660–675.

Cohen, A. 1983. Comparing regression coefficients across subsamples: A study of the statistical test. *Sociological Methods & Research* 12 (1): 77-94.

Crosetto, P., and A. Filippin. 2013. The "bomb" risk elicitation task. *Journal of Risk and Uncertainty* 47 (1): 31–65.

Dittmann, I., E. Maug, and D. Zhang. 2011. Restricting CEO pay. *Journal of Corporate Finance* 17 (4): 1200–1220.

Directive 2013/36/EU. CRD IV. European Parliament and the Council. *Official Journal of the European Union*. June 26. Available at: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:176:0338:0436:En:PDF.

Farrell, A. M., J. H. Grenier, and J. Leiby. 2017. Scoundrels or Stars? Theory and Evidence on the Quality of Workers in Online Labor Markets. *The Accounting Review* 92 (1): 93–114.

Festinger, L. 1957. *A Theory of Cognitive Dissonance.* Stanford, Calif.: Stanford Univ. Press.

Garner, J. L., and W. Y. Kim. 2010. Does a Salary Cap Really Work? Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1546783.

Grenier, J. H., B. Pomeroy, and M. T. Stern. 2015. The Effects of Accounting Standard Precision, Auditor Task Expertise, and Judgment Frameworks on Audit Firm Litigation Exposure. *Contemporary Accounting Research* 32 (1): 336–357.

Hartmann, F., and S. Slapničar. 2015. An experimental study of the effects of negative, capped and deferred bonuses on risk taking in a multi-period setting. *Journal of Management & Governance* 19 (4): 875–896.

Hendriks, A. 2012. SoPHIE - Software Platform for Human Interaction Experiments. *Working Paper*.

Holland, R. W., R. M. Meertens, and M. van Vugt. 2002. Dissonance on the Road. Self-Esteem as a Moderator of Internal and External Self-Justification Strategies. *Personality and Social Psychology Bulletin* 28 (12): 1713–1724.

Horton, J. J., and L. B. Chilton. 2010. *The Labor Economics of Paid Crowdsourcing.* New York, NY: Association for Computing Machinery.

Jansen, E. P., K. A. Merchant, and W. A. van der Stede. 2009. National differences in incentive compensation practices. The differing roles of financial performance measurement in the United States and the Netherlands. *Accounting, Organizations and Society* 34 (1): 58–84.

Jensen, M. C., and W. H. Meckling. 1976. Theory of the Firm. Managerial Behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics* 3: 305–360.

Jokivuolle, E., J. Keppo, and X. Yuan. 2015. Bonus Caps, Deferrals and Bankers' Risk-Taking. *Bank of Finland Research Discussion Paper* (5): 1–75.

Kahneman, D., and A. Tversky. 1979. Prospect Theory. An Analysis of Decision under Risk. *Econometrica* 47 (2): 263.

Kahneman, D., and A. Tversky. 1981. The Psychology of Preferences. *Scientific American* 246 (1): 160–173.

Kleymenova, A., and A. I. Tuna. 2016. Regulation of Compensation. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2755621.

Koonce, L., J. Miller, and J. Winchel. 2015. The Effects of Norms on Investor Reactions to Derivative Use. *Contemporary Accounting Research* 32 (4): 1529–1554.

Lefebvre, M., and F. M. Vieider. 2013. Reining in excessive risk-taking by executives. The effect of accountability. *Theory and Decision* 75 (4): 497–517.

Lerner, J. S., and P. E. Tetlock. 1999. Accounting for the effects of accountability. *Psychological Bulletin* 125 (2): 255–275.

Libby, T., S. E. Salterio, and A. Webb. 2004. The Balanced Scorecard. The Effects of Assurance and Process Accountability on Managerial Judgment. *The Accounting Review* 79 (4): 1075–1094.

Lönnqvist, J.-E., M. Verkasalo, G. Walkowitz, and P. C. Wichardt. 2015. Measuring individual risk attitudes in the lab. Task or ask? An empirical comparison. *Journal of Economic Behavior & Organization* 119: 254–266.

Low, A. 2009. Managerial risk-taking behavior and equity-based compensation. *Journal of Financial Economics* 92 (3): 470–490.

Luce, R. D. 1977. The choice axiom after twenty years. *Journal of Mathematical Psychology* 15 (3): 215–233.

Luce, R. D. 2005. *Individual choice behavior: A theoretical analysis.* Mineola, N.Y: Dover Publications.

Meade, A. W., and S. B. Craig. 2012. Identifying careless responses in survey data. *Psychological methods* 17 (3): 437–455.

Murphy, K. J. 2001. Performance Standards in Incentive Contracts. *Journal of Accounting and Economics* 30: 245–278.

Murphy, K. J. 2013. Regulating Banking Bonuses in the European Union. A Case Study in Unintended Consequences. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2235395.

Neumann, J. von, and O. Morgenstern. 2007. *Theory of games and economic behavior.* 60th anniversary ed. Princeton classic editions. Princeton: Princeton University Press.

Neumann, N., U. Böckenholt, and A. Sinha. 2016. A meta-analysis of extremeness aversion. *Journal of Consumer Psychology* 26 (2): 193–212.

Pahlke, J., S. Strasser, and F. M. Vieider. 2012. Risk-taking for others under accountability. *Economics Letters* 114 (1): 102–105.

Paolacci, G., J. Chandler, and P. G. Ipeirotis. 2010. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making* 5 (5): 411–419.

Pollmann, M. M.H., J. Potters, and S. T. Trautmann. 2014. Risk taking by agents. The role of ex-ante and ex-post accountability. *Economics Letters* 123 (3): 387–390.

Rennekamp, K. 2012. Processing Fluency and Investors' Reactions to Disclosure Readability. *Journal of Accounting Research* 50 (5): 1319–1354.

Samuelson, W., and R. Zeckhauser. 1988. Status quo bias in decision making. *Journal of Risk and Uncertainty* 1 (1): 7–59.

Schedlinsky, I., F. Sommer, and A. Wöhrmann. 2018. Influencing Risk Taking in Competitive Environments: An Experimental Analysis. *Journal of Risk Finance* (forthcoming).

Simonson, I. 1989. Choice Based on Reasons: The Case of Attraction and Compromise Effects. Journal of Consumer Research. *Journal of Consumer Research* 16 (2): 158–174.

Simonson, I. 2014. Vices and Virtues of Misguided Replications. The Case of Asymmetric Dominance. *Journal of Marketing Research* 51 (4): 514–519.

Sprinkle, G. B., M. G. Williamson, and D. R. Upton. 2008. The effort and risk-taking effects of budget-based contracts. *Accounting, Organizations and Society* 33 (4-5): 436–452.

Tetlock, P. E. 1985. Accountability. The neglected social context of judgment and choice. *Research in Organizational Behavior* (7): 297–332.

Tetlock, P. E., F. M. Vieider, S. V. Patil, and A. M. Grant. 2013. Accountability and ideology. When left looks right and right looks left. *Organizational Behavior and Human Decision Processes* 122 (1): 22–35.

Tversky, A., and D. Kahneman. 1991. Loss Aversion in Riskless Choice. A Reference-Dependent Model. *The Quarterly Journal of Economics* 106 (4): 1039–1061.

Tversky, A., and I. Simonson. 1992. Choice in Context: Tradeoff Contrast and Extremeness Aversion. *Journal of Marketing Research* 29 (3): 281–295.

Tversky, A., and I. Simonson. 1993. Context-Dependent Preferences. *Management Science* 39 (10): 1179–1189.

Vieider, F. M. 2009. The effect of accountability on loss aversion. *Acta psychologica* 132 (1): 96–101.

Weigold, M. F., and B. R. Schlenker. 1991. Accountability and Risk Taking. *Personality and Social Psychology Bulletin*, 17 (1): 25–29.

Table 1

Descriptive statistics (mean, [standard deviation])

	Cap[a]						**Total**
	No			Yes			
	Justification[b]			Justification[b]			
	Low	High	**Total**	Low	High	**Total**	
All subjects[c]							
Number of subjects	114	119	**233**	108	106	**214**	**447**
Number of boxes collected[d]	32.39	36.63	**34.56**	30.70	30.75	**30.73**	**32.72**
	[13.93]	[15.46]	**[14.85]**	[11.21]	[14.10]	**[12.70]**	**13.98**
Only risk-averse subjects[c]							
Number of subjects	67	68	**135**	76	64	**140**	**275**
Number of boxes collected[d]	27.99	28.90	**28.44**	27.92	24.14	**26.19**	**27.30**
	[10.34]	[9.15]	**[9.73]**	[9.33]	[10.45]	**[10.00]**	**[9.92]**

[a] *Cap* is manipulated between subjects at two levels. In the no-cap condition, participants choose to collect between 0 and 100 boxes (BRET task, Crosetto and Filippin 2013) and receive \$0.03 per box if the bomb is not collected. In the cap condition, participants also choose to collect between 0 and 100 boxes and receive \$0.03 per box if the bomb is not collected. However, compensation in the cap condition is limited to \$1.5, equal to the compensation for collecting 50 boxes.

[b] *Justification* is manipulated between subjects at two levels. In the low-justification condition, participants can proceed after the BRET task without any need for external justification. Before the BRET task starts, participants in the high-justification condition receive information that they have to give a written justification about their decision.

[c] *Risk preference* is measured with a risk-elicitation instrument (Sprinkle et al. 2008). This instrument requires participants to make fifteen choices between receiving a certain amount of \$0.75 and participating in a lottery. The lotteries consist of a chance of winning \$1.50 with a probability of π and \$0 with a probability of $1-\pi$ (π varies from 85% to 15% in 5% increments). Once participants make their fifteen choices, one of the fifteen choices is randomly selected, and participants' earnings are determined based on their choice and the result of the lottery. Participants are declared risk-averse if they switch prior to the eighth alternative and select fewer than 50 boxes.

[d] *Selected boxes* represents the number of boxes participants collect in the BRET task.

Table 2

Test H1: Results of capped compensation and justification for risk-taking for all subjects (ANOVA)

Dependent variable: Number of boxes collected (n = 447)

Source	Df	MS	F-Statistic	p-value
Cap	1	1,596.19	8.37	< 0.01[a]
Justification	1	512.27	2.69	0.10[b]
Cap x Justification	1	488.17	2.56	0.11[b]

[a] The p-value is reported on a one-tailed basis due to the directional hypothesis for this effect.

[b] The p-value is reported on a two-tailed basis due to the lack of a directional hypothesis for this effect.

Table 3

Test of H2: Results of capped compensation and justification on risk-taking for risk-averse subjects

Panel A: ANOVA

Dependent variable: Number of boxes collected (n = 275)

Source	Df	MS	F-Statistic	p-value
Cap	1	397.79	4.13	0.02[a]
Justification	1	140.86	1.46	0.23[b]
Cap x Justification	1	376.94	3.92	0.02[a]

Panel B: Model contrast[c]

Dependent variable: Number of boxes collected (n = 275)

Source	Df	MS	F-Statistic	p-value
Model contrast	1	682.40	7.09	< 0.01[a]

[a] The p-value is reported on a one-tailed basis due to the directional hypothesis for this effect.

[b] The p-value is reported on a two-tailed basis due to the lack of a directional hypothesis for this effect.

[c] The contrast coefficients are +1 for no cap/low justification, +3 for no cap/high justification, -1 for cap/low justification, and -3 for cap/high justification.

Table 4

Descriptive statistics for risk-averse subjects with and without target (mean, [standard deviation])

	Cap[a]						**Total**
	No Justification[b]			Yes Justification[b]			
	Low	High	**Total**	Low	High	**Total**	
Only risk-averse subjects[c]							
Number of subjects	67	68	**135**	76	64	**140**	**275**
Number of boxes collected[d]	27.99	28.90	**28.44**	27.92	24.14	**26.19**	**27.30**
	[10.34]	[9.15]	**[9.73]**	[9.33]	[10.45]	**[10.00]**	**[9.92]**
Risk-averse subjects without target[c,e]							
Number of subjects	29	26	**55**	32	33	**65**	**120**
Number of boxes collected[d]	25.93	29.15	**27.45**	24.53	21.15	**22.82**	**24.94**
	[10.11]	[9.93]	**[10.07]**	[9.64]	[9.43]	**[9.61]**	**[10.05]**

[a] *Cap* is manipulated between subjects at two levels. In the no-cap condition, participants choose to collect between 0 and 100 boxes (BRET task, Crosetto and Filippin 2013) and receive \$0.03 per box if the bomb is not collected. In the cap condition, participants also choose to collect between 0 and 100 boxes and receive \$0.03 per box if the bomb is not collected. However, compensation in the cap condition is limited to \$1.5, equal to the compensation for collecting 50 boxes.

[b] *Justification* is manipulated between subjects at two levels. In the low-justification condition, participants can proceed after the BRET task without any need for external justification. Before the BRET task starts, participants in the high-justification condition receive information that they have to give a written justification about their decision.

[c] *Risk preference* is measured with a risk-elicitation instrument (Sprinkle et al. 2008). This instrument requires participants to make fifteen choices between receiving a certain amount of \$0.75 and participating in a lottery. The lotteries consist of a chance of winning \$1.50 with a probability of π and \$0 with a probability of $1-\pi$ (π varies from 85% to 15% in 5% increments). Once participants make their fifteen choices, one of the fifteen choices is randomly selected, and participants' earnings are determined based on their choice and the result of the lottery. Participants are declared risk-averse if they switch prior to the eighth alternative and select fewer than 50 boxes.

[d] *Selected boxes* represents the number of boxes participants collect in the BRET task.

[e] *Risk-averse subjects without target* include only participants who stated "No" when asked "Did you aim at earning a specific compensation from the bomb task?"

Table 5

Test of H2 replication: Results of capped compensation and justification for risk-taking for risk-averse subjects without target[d]

Panel A: ANOVA

Dependent variable: Number of boxes collected (n = 275)

Source	Df	MS	F-Statistic	p-value
Cap	1	3,842.40	21.07	< 0.01[a]
Justification	1	12.00	0.07	0.80[b]
Cap x Justification	1	530.96	2.91	0.05[a]

Panel B: Model contrast[c]

Dependent variable: Number of boxes collected (n = 275)

Source	Df	MS	F-Statistic	p-value
Model contrast	1	4,325.70	23.72	< 0.01[a]

[a] The p-value is reported on a one-tailed basis due to the directional hypothesis for this effect.

[b] The p-value is reported on a two-tailed basis due to the lack of a directional hypothesis for this effect.

[c] The contrast coefficients are +1 for no cap/low justification, +3 for no cap/high justification, -1 for cap/low justification, and -3 for cap/high justification.

[d] Risk-averse subjects without a target include only participants who stated “No” when we asked them “Did you aim at earning a specific compensation from the bomb task?”

Figure 1

Predicted interaction effect of capped compensation and justification on risk-taking for risk-averse subjects (H2b)

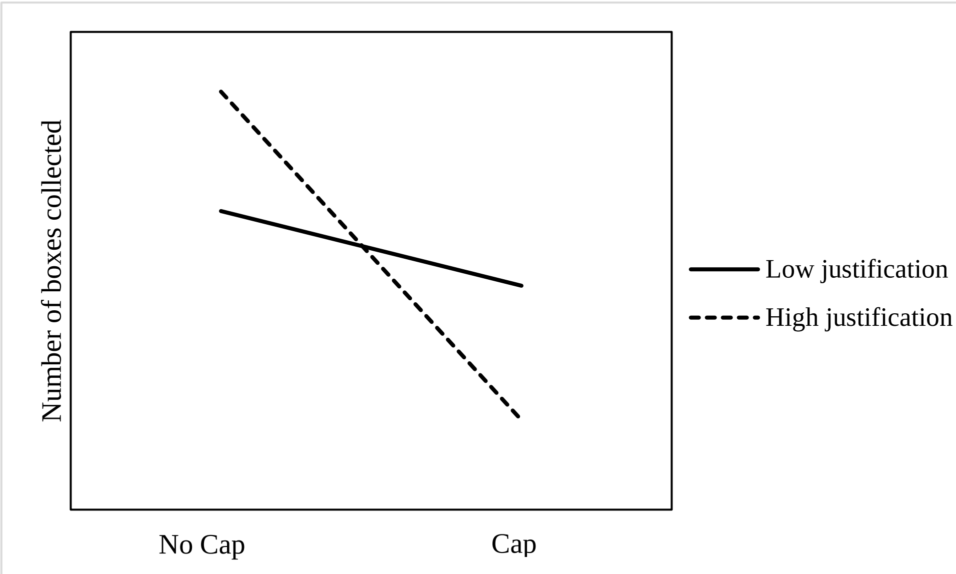

Figure 2

BRET Task

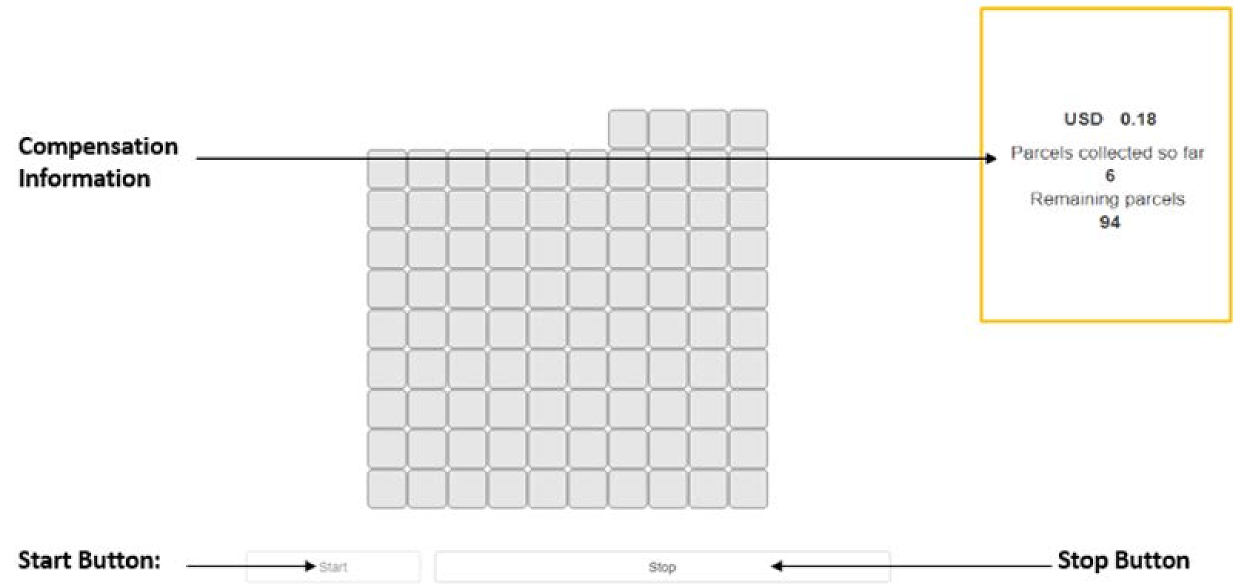

Figure 3

Experimental Procedure

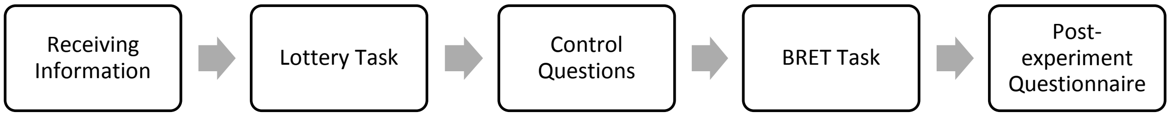

Figure 4

Observed interaction effect of capped compensation and justification on risk-taking for risk-averse subjects (H2b)

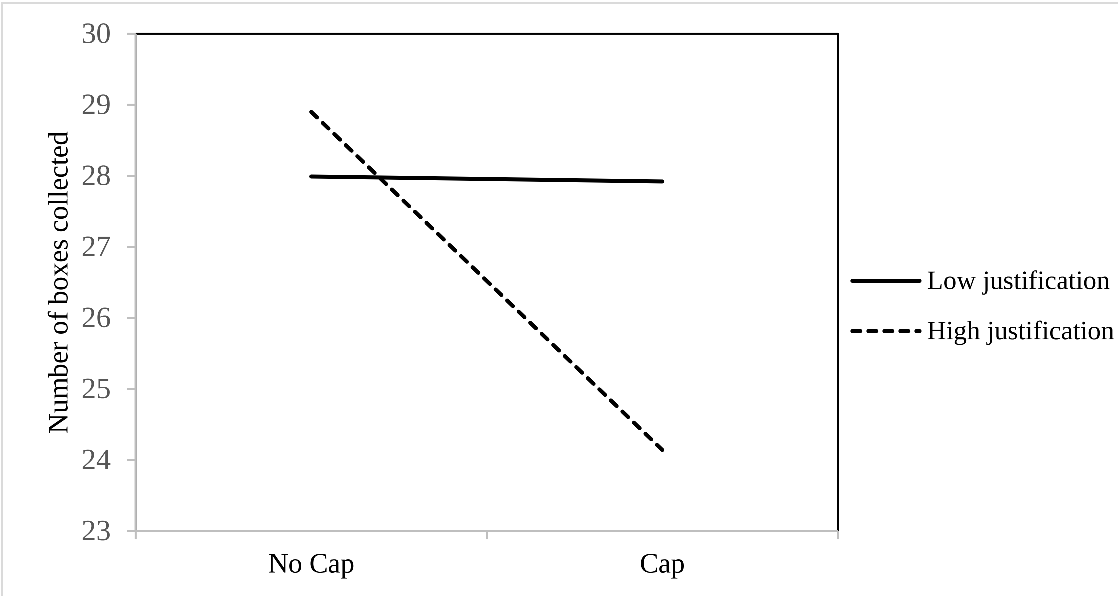

Figure 5

Observed interaction effect of capped compensation and justification on risk-taking for risk-averse subjects without target (H2b replication)