European Economic Review 187 (2026) 105363

Contents lists available at ScienceDirect

European Economic Review

journal homepage: www.elsevier.com/locate/eer

Performance thresholds, effort and risk-taking

Pol Campos-Mercade [a,b], Sergio Currarini [c,d,*], Carmen Marchiori [e], Friederike Mengel [f,g]

[a] Lund University, Sweden
[b] Institute for Future Studies, Sweden
[c] University of Leicester, United Kingdom
[d] Universita' Ca Foscari di Venezia, Italy
[e] Universita' degli Studi di Brescia, Italy
[f] Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK
[g] Department of Economics, Erasmus University Rotterdam, The Netherlands

ARTICLE INFO

Keywords:
Performance-based thresholds
Effort
Risk-taking

ABSTRACT

We study, theoretically and experimentally, the impact of performance thresholds on effort and risk-taking. Our theory predicts that thresholds typically increase effort provision, but that the magnitude of this increase varies *non-monotonically* with the threshold level. For risk-taking, the effects are positive for high thresholds and negative for low thresholds. We characterize these patterns through the subtle interplay between effort and risk, which transition from *complements* to *substitutes* as thresholds become more demanding. Our experiment (N=1092) supports the theoretical predictions, showing that high thresholds raise both effort and risk-taking, whereas more moderate thresholds raise effort but lower risk-taking. Importantly, the experiment sheds light on the crucial role of individual heterogeneity: because individuals have different baseline performances, the same threshold can be perceived as high by some and low by others, producing opposite behavioral responses. The analysis highlights that effective threshold design must understand not only how thresholds shape effort and risk-taking, but also the characteristics of the targeted population.

1. Introduction

Performance thresholds are ubiquitous across modern economic and organizational life. Students face minimum grades to pass exams. Researchers must meet publication standards for tenure and promotion. Mutual fund managers earn bonuses only if their portfolios outperform a target return. Athletes face performance thresholds to qualify for the Olympics and other major competitions. Countries are subject to fulfilling certain standards to be admitted to an economic union.

These thresholds are typically designed to induce higher effort. Yet in all these settings, agents decide not only how much effort to exert, but also how much risk to take. For example, a student may decide to study the entire syllabus or gamble on a subset of topics; a researcher may pursue several safe projects or concentrate all her effort on a single ambitious idea; a fund manager may invest in a well-diversified portfolio or speculate in high-risk, high-return assets. Understanding how performance thresholds shape the joint choice of effort and risk-taking is therefore crucial for assessing their economic implications and designing effective incentive schemes.

* Corresponding author at: University of Leicester, United Kingdom.
E-mail addresses: pol.campos@nek.lu.se (P. Campos-Mercade), sc526@le.ac.uk (S. Currarini), carmen.marchiori@unibs.it (C. Marchiori), fr.mengel@gmail.com (F. Mengel).

https://doi.org/10.1016/j.euroecorev.2026.105363
Received 14 June 2025; Received in revised form 28 April 2026; Accepted 2 May 2026
Available online 13 May 2026

In this paper, we study how the discontinuities in payoff structures induced by performance thresholds affect these two fundamental dimensions of choice. We do so by combining a stylized theoretical model with a pre-registered experiment involving 1092 participants.

In our model, an agent exerts costly effort to produce an output from which she derives utility. We assume that effort can be allocated across two production technologies: a *riskless* one, converting each unit of effort into one unit of output; and a *risky* technology yielding, for each unit of effort, either a high output or a low output, with equal probability. The agent therefore faces two decisions: how much effort to exert, and what fraction of that effort to allocate to the risky technology. After output is realized, the agent receives a bonus if output exceeds some threshold reference level and incurs a loss otherwise.

To make the theoretical structure more concrete, consider the example of a student preparing for an exam. Here, effort corresponds to the amount of time she devotes to studying, while risk captures how she chooses to study—whether she prepares the entire syllabus (a safer strategy) or focuses her study time on only a few topics (a riskier strategy). Concentrating the same amount of study time on fewer topics makes her much better prepared on those topics than she would be if she had distributed her time across the entire syllabus. As a result, if one of the selected topics appears on the exam, she attains a higher mark than under the safer strategy; if none of them appears, her performance is substantially worse and she may fail the exam. The threshold corresponds to the minimum grade required to pass.

Our experimental design mirrors the theoretical framework closely: participants make choices about effort and risk both in settings without thresholds and in settings where they receive additional payment for reaching a specific threshold. They provide effort in a real effort task, specifically by counting the number of zeroes in binary matrices (Abeler et al., 2011), and take risk by choosing lotteries over returns to effort.

Our analysis delivers two main take-home messages. First, the interplay of effort and risk-taking in shaping behavior depends on the degree of "compliance" with the threshold, and can be understood in terms of the familiar notions of complementarity and substitution. When agents meet the threshold in all states ("full compliance"), effort and risk are *complements*: taking more risk requires more effort to ensure that the threshold is met in the bad state. This makes agents more cautious, in the attempt to limit the use of costly effort. This reduction in risk-taking is not due to risk aversion, and would hold for risk-neutral agents, even in a more general statistical structure with full support, as we show in the appendix. When agents meet the threshold only in the good state ("partial compliance"), effort and risk are *substitutes*: taking more risk reduces the effort needed to reach the threshold. This makes agents less cautious.

Second, what type of interplay is operative depends on how demanding the threshold is relative to the agent's "status quo" performance. Meeting an increasingly demanding threshold requires increasingly costly adjustments compared to the status quo. As a result, optimal behavior transitions from full to partial compliance (and, eventually, to no compliance) as the threshold increases. This natural transition yields *non-monotonic* patterns in both effort and risk-taking. More precisely, while at first more stringent thresholds prescribe an increase in effort and a reduction in risk as the optimal way to fully comply, once agents switch from full to partial compliance (in response to more demanding thresholds), effort drops sharply and risk-taking rises (acting as a substitute for costly effort). Despite this sudden drop, effort remains above the "status quo" level for all thresholds (except for very low ones, for which no adjustment is required). The effect on risk-taking crucially depends, instead, on the threshold level: risk-taking decreases in response to moderate (though meaningful) thresholds, while it increases for very low and very high thresholds.

To illustrate these mechanisms, consider again the example of a student preparing for an exam. When the pass mark is not too high relative to the student's baseline performance, she aims to pass with certainty. In this case (full compliance), taking risk by studying only part of the syllabus would require additional compensating effort to ensure passing even in the "bad" realization (complementarity of effort and risk). Hence selective studying is not very attractive. In contrast, if the pass mark is very demanding, the student may find it overly challenging to guarantee passing. In this case (partial compliance), studying selectively becomes attractive because additional risk substitutes for costly effort.

Our experiment provides systematic evidence in support of these mechanisms and highlights the key role of heterogeneity. Consistent with the theoretical predictions for partial compliance, participants on average exert more effort and take more risk when they face relatively high thresholds. By contrast, we observe little to no effect on both effort and risk-taking when a low threshold is introduced. To make sense of these average effects, we account for individual heterogeneity: since participants differ in their baseline performance, the same threshold can be "high" for some and "low" for others. When we sort participants by their status quo choices, we find that the same threshold induces very different reactions from different groups of subjects. This pattern is not noise: it is precisely what the model predicts when individuals differ in their distance from the threshold under the status quo. In particular, those who "perceive" the threshold as high (relative to their status quo position) increase both effort and risk, whereas those for whom the threshold is relatively low tend to increase effort and reduce risk-taking.

Our results highlight the potential but also the pitfalls of using thresholds to create incentives. When the goal is to minimize risk-taking – for instance, in educational settings where excessive risk may lead students to neglect part of the material – thresholds should be set not too high so that full compliance is optimal, making effort and risk complements. Conversely, when risk-taking is desirable – as in academic research or entrepreneurial environments where exploration is valuable – thresholds should be sufficiently demanding to induce partial compliance, thereby increasing risk-taking through the substitution effect. Importantly, because individuals differ in their baseline performance, the same threshold may generate opposite behavioral responses across agents. Effective threshold design must therefore take into account not only the intended direction of responses, but also the heterogeneity of the targeted population.

The paper is organized as follows. Section 2 reviews the related literature and explains how our paper contributes to existing research. Sections 3 and 4 set out the model with and without a threshold, and characterize optimal solutions. Section 5 presents the experimental design and results. Section 6 concludes.

2. Related literature

There are a number of different literatures that our research contributes to. First, threshold-based compensation schemes have been studied in a number of applied settings including education and health. In education much of the research has focused on studying how students' performance is affected by standards set for attaining scholarships (Cornwell et al., 2005; Scott-Clayton, 2011; Cohodes and Goodman, 2014; Bettinger et al., 2019; Fack and Grenet, 2015; Montalbán, 2018) and monetary rewards (Kremer et al., 2009; Leuven et al., 2010; De Paola et al., 2012; Levitt et al., 2017; Campos-Mercade and Wengström, 2020; Campos-Mercade et al., 2025). This literature often shows positive effects on performance for students around the threshold, although it provides limited understanding of the extent to which incentives affect effort (such as studying more) and risk-taking (such as studying parts of the material differently). In health economics, thresholds have been studied because of the increasingly widespread adoption of incentive schemes based on performance targets, whose attainment triggers a flat financial bonus (Forsberg et al., 2001; Volpp et al., 2008; Campbell et al., 2009; Charness and Gneezy, 2009; Gravelle et al., 2010; Lacetera et al., 2013; Milkman et al., 2021; Campos-Mercade et al., 2021; 2024b). Much of this literature has focused on the impact on the effort exerted by physicians. Our work provides novel insights into how these widespread incentives affect individuals' decision-making in terms of effort provision as well as risk-taking.

Second, there is also an extensive theoretical literature on performance-based compensation for managers and executives. This literature has framed the problem as a principal–agent relation (Jensen and Meckling, 1976), where managers are incentivized to exert effort by means of performance-based compensation schemes. In response to such payments, risk-averse managers may reduce personal risk below the firm's optimal level (Miller et al., 2002; Demski and Feltham, 1978; Amihud and Lev, 1981; Smith and Stulz, 1985). Stock option-based compensation schemes have been proposed as a way to incentivize risk-taking by convexifying the manager's payoff function (Smith and Stulz, 1985). Compensation schemes based on stock options can also be viewed as incentive contracts based on performance thresholds, which trigger a bonus if the threshold is met (Lambert and Larcker, 1991; Murphy, 2001).[1] Although our problem is not framed as a principal–agent problem, our solution with a threshold can still be thought of as characterizing the set of incentive-compatible outcomes, given that the contract proposed by the principal is made of a threshold. Such contracts are not studied in the existing literature on performance-based compensation, with the exception of Palomino and Prat (2003), who show that a flat bonus (and no penalty) contract can be optimal with limited liability. Zhou and Swan (2003) also study discontinuous contracts, and show that these can be optimal in a principal–agent model where the only unobserved choice is effort. It is important to note that, by studying a two-state model, where the threshold can be met with certainty, we are able to study the interplay between effort and risk-taking purely as ways to meet the threshold in both states, in one state, or never. Bonus contracts are also studied in the field of accounting (budget-based contracts). Sprinkel et al. (2008) design an experiment to study the interplay of effort and risk when effort is rewarded by means of a performance-based threshold. They find that risk-taking can be used as a way to attain a very high threshold, which cannot be attained through effort alone. A similar role for risk is found, within the field of finance, by Shelev (2012), where high risk is used only when the threshold is set at high enough levels. These findings are consistent with our results under "partial compliance", where risk is used as a substitute for effort.

Third, our paper also speaks to a theoretical and experimental literature that studies individuals' behavior in contests (for reviews, see Konrad et al., 2009; Dechenaux et al., 2015).[2] Contests in which a set of winners gain a prize provide a similar setting to threshold incentives, the main difference being that the threshold is not fixed but rather endogenous to the agents' decisions. Hence, our results also speak to contests in which agents can more or less accurately guess where the threshold will be (e.g., repeated contests with many participants, such as university admission cutoffs, or situations in which agents know or correctly guess the others' decisions). First, we find that participants start taking considerable risk when the threshold reaches high levels, and reduce effort at the same time. This finding echoes the result in the contest literature that risk and effort are substitutes under many conditions (Hvide, 2002; Gilpatric, 2009; Nieken, 2010; Andersson et al., 2020; although see Gilpatric, 2009; Andersson et al., 2020 for settings in which risk and effort could be complements). Second, our finding that participants are highly heterogeneous likely extends to contests, indicating that it is essential to account for heterogeneity to model these situations.

Fourth, our work also relates to an experimental literature in psychology and economics studying how reference points affect effort provision. This literature finds that thresholds, benchmarks, and goals often act as reference points consistent with (Kahneman and Tversky, 1979)'s Prospect Theory (e.g., Heath et al., 1999; Abeler et al., 2011; Gill and Prowse, 2012; Ockenfels et al., 2014; Campos-Mercade et al., 2024a), and they typically increase effort provision (Locke and Latham, 1990; Camerer et al., 1997; Goerg and Kube, 2012; Corgnet et al., 2015, 2018; González-Jiménez, 2024; Koch and Nafziger, 2020). However, this literature has focused on studying the effects of thresholds on effort, without allowing agents the possibility to take risks, as it is common in reality. We complement this literature by studying the effects of thresholds when people can choose not only their effort provision, but both effort and risk. Since thresholds can be attained not only through higher effort but also by taking risk, our paper provides a more complete picture of people's decision-making in such situations.

[1] In stock option contracts, the exercise price plays the role of the threshold, defined on the stock value which is under the control of the executives.

[2] Most studies in the literature on contests and competition have focused on studying either effort provision (see e.g. Tullock, 1980; Lazear and Rosen, 1981; Nalebuff and Stiglitz, 1983) or risk-taking (see e.g. Dekel and Scotchmer, 1999; Tsetlin et al., 2004; Eriksen and Kvaløy, 2017). However, a few theoretical (Hvide, 2002; Gilpatric, 2009) and experimental (Kräkel et al., 2014; Nieken, 2010; Andersson et al., 2020) papers study both effort and risk-taking, often finding that competition increases both dimensions. Interestingly, Eriksen and Kvaløy (2017) find that competition per se triggers risk-taking, even when the optimal decision is not to take any risk.

Finally, our modeling set-up is formally equivalent to a two-period, two-asset and two-state saving and portfolio model (Sandmo, 1968, 1970; Aura et al., 2002; Gollier, 2008); our analysis enriches the basic model by imposing a threshold on the ex-post realization of the portfolio, with gains and losses accruing accordingly (see Section 3.1 and Appendix A.1). Our results provide insights on the effect of such thresholds on savings and on the risky component of the optimal portfolio.

3. Model

An agent exerts effort e in the production of a good y from which she draws utility $v(y)$. For the student preparing for an exam, e is the amount of time spent studying for an exam and y is the mark obtained in the exam. Effort is costly, and the cost of effort is represented by the strictly convex cost function $C(e)$. The agent's total utility from the pair (e, y) is therefore measured by:

$$U(e, y) = v(y) - C(e),$$

where v is assumed strictly increasing, concave and three times differentiable. We also assume that $C'(0) = 0$ and that $\lim_{e\to\infty} C'(e) = \infty$. Two production technologies transform effort into good y. The first is riskless: e units of effort are transformed into e units of y. The second is risky and given by:

$$\tilde{y} = \begin{cases} e(1+R) & \text{with prob. } \frac{1}{2} \\ e(1-r) & \text{with prob. } \frac{1}{2} \end{cases} \tag{1}$$

where $R \geq 1 > r > 0$. We allow for technology mixes and denote by $\alpha \in [0, 1]$ the fraction of effort invested in the risky technology. In the example of the student, α measures the extent to which the student focuses only on specific topics in their preparation or prepares the entire syllabus. The random level of output generated by the pair (α, e) is therefore:

$$\tilde{y}(\alpha, e) = \begin{cases} y_H(\alpha, e) \equiv e(1+\alpha R) & \text{with prob. } \frac{1}{2} \\ y_L(\alpha, e) \equiv e(1-\alpha r) & \text{with prob. } \frac{1}{2}. \end{cases} \tag{2}$$

The ex-post realization of the random variable $\tilde{y}(e, \alpha)$ is then compared to an exogenous threshold Y to determine the final level of consumption $\tilde{y}(e, \alpha, Y)$ of commodity y. In particular:

- if $\tilde{y}(\alpha, e) \geq Y$ then the consumption of good y is augmented by the amount $G \geq 0$ (bonus or "gain")

$$\tilde{y}(\alpha, e, Y) = \tilde{y}(e, \alpha) + G \tag{3}$$

- if $\tilde{y}(\alpha, e) < Y$ then the consumption of good y is reduced by the amount $L \geq 0$ (penalty or "loss").

$$\tilde{y}(\alpha, e, Y) = \tilde{y}(\alpha, e) - L \tag{4}$$

We identify three possible regions associated with the pair (e, α):

$$\Omega_1 = \left\{(\alpha, e) : y_H(\alpha, e) < Y\right\}; \tag{5}$$
$$\Omega_2 = \left\{(\alpha, e) : y_L(\alpha, e) < Y \leq y_H(\alpha, e)\right\}; \tag{6}$$
$$\Omega_3 = \left\{(\alpha, e) : y_L(\alpha, e) \geq Y\right\} \tag{7}$$

In region Ω_1 the threshold is never met; in region Ω_2 the threshold is met in the good state of the world only; in region Ω_3 the threshold is met in both states of the world. Such regions identify different levels of "compliance" with the threshold. We will refer to region Ω_2 as *partial compliance* and to region Ω_3 as *full compliance*. We will also refer to the interior and the boundary of regions Ω_2 and Ω_3 with the following notation:

$$\bar{\Omega}_2 = \left\{(\alpha, e) : y_H(\alpha, e) = Y\right\}; \tag{8}$$
$$\Omega_2^o = \left\{(\alpha, e) : y_H(\alpha, e) > Y, y_L(\alpha, e) < Y\right\}; \tag{9}$$
$$\bar{\Omega}_3 = \left\{(\alpha, e) : y_L(\alpha, e) = Y\right\}. \tag{10}$$
$$\Omega_3^o = \left\{(\alpha, e) : y_L(\alpha, e) > Y\right\}. \tag{11}$$

We define the following expected utilities, associated with regions 1, 2 and 3, respectively:

$$E_1\tilde{U}(\alpha, e) = \left[\frac{1}{2}v\left(y_L(\alpha, e) - L\right) + \frac{1}{2}v\left(y_H(\alpha, e) - L\right)\right] - C(e) \tag{12}$$
$$E_2\tilde{U}(\alpha, e) = \left[\frac{1}{2}v\left(y_L(\alpha, e) - L\right) + \frac{1}{2}v\left(y_H(\alpha, e) + G\right)\right] - C(e) \tag{13}$$
$$E_3\tilde{U}(\alpha, e) = \left[\frac{1}{2}v\left(y_L(\alpha, e) + G\right) + \frac{1}{2}v\left(y_H(\alpha, e) + G\right)\right] - C(e) \tag{14}$$

Overall expected utility is therefore defined as:

$$E\tilde{U}(\alpha, e) = \begin{cases} E_1\tilde{U}(\alpha, e) \;\; if \;\; (e, \alpha) \in \Omega_1 \\ E_2\tilde{U}(\alpha, e) \;\; if \;\; (e, \alpha) \in \Omega_2 \\ E_3\tilde{U}(\alpha, e) \;\; if \;\; (e, \alpha) \in \Omega_3 \end{cases}$$

It should be noted that the introduction of a threshold, and of the relative gains and losses, is by no means equivalent to a spread of the returns of the risk asset, whose comparative statics is studied, for instance, in Gollier (2008). This is immediately clear when we consider regions Ω_1 and Ω_3, where either the loss or the bonus applies in both states of the world. Within region Ω_2, where the threshold is met only in the good state, there is an exogenous spread in the returns of the whole portfolio due to the gains and losses associated with the good and bad state of the world, respectively. However, as we show in Appendix A.1, this cannot be expressed as an exogenous spread in the returns of the risky asset; rather it is an endogenous spread, whose magnitude depends on both choice variables e and α.

We now formulate the decision problem with a threshold. Given Y, the agent solves the following maximization problem:

$$\max_{(\alpha,e)} E\tilde{U}(\alpha,e) \tag{15}$$
$$s.t.\ \ 0 \le \alpha \le 1; e \ge 0.$$

This problem is discontinuous at the frontier of the regions defined in (5)–(7). Hence, we first need to show that it admits a solution, as stated in the next proposition.

Proposition 1. *Problem (15) admits a solution.*

A solution to (15) can be viewed as a two-stage process. First, for the three possible compliance regions, the best choice of effort and risk-taking is selected and the relative (expected) welfare is evaluated. Second, these maximal levels of welfare are compared to select the best plan. Necessary conditions for a best choice of effort and risk-taking for each compliance region are provided in the Appendix (Lemmas 2 to 6).

The case $G = L = 0$ is equivalent to a decision problem without any threshold. As such, it can be thought of as a "status quo" situation. This case is of interest if one wishes to evaluate the very effect of introducing a threshold in an otherwise threshold-free problem. We will denote the status quo optimal choice as (α^*, e^*). Standard arguments imply that $\alpha^* > 0$, as long as v is strictly concave and $R > r$. Moreover, the conditions imposed on C imply that $\infty > e^* > 0$. We will further assume that $\alpha^* < 1$.[3] Under these specifications, the following first-order conditions characterize the unique status quo interior optimum:

$$\frac{\partial}{\partial e} = 0 \to \frac{1}{2}v'_H \cdot (1+\alpha^* R) + \frac{1}{2}v'_L \cdot (1-\alpha^* r) = C'(e); \tag{16}$$
$$\frac{\partial}{\partial \alpha} = 0 \to \frac{1}{2}v'_H R = \frac{1}{2}v'_L r, \tag{17}$$

where we have used the following notation:

$$v'_H \equiv v'(y_H(e^*,\alpha^*));\ \ v'_L \equiv v'(y_L(e^*,\alpha^*)). \tag{18}$$

Substituting (17) into (16) we obtain:

$$v'(e^*(1-\alpha^* r)) = C'(e^*)\frac{2R}{(r+R)}; \tag{19}$$
$$v'(e^*(1+\alpha^* R)) = C'(e^*)\frac{2r}{(r+R)}. \tag{20}$$

To ease notation, we will write $y^*_L \equiv y_L(e^*,\alpha^*)$ and $y^*_H \equiv y_H(e^*,\alpha^*)$.

4. Theoretical results

How does the introduction of a threshold affect behavior? Do more demanding thresholds always incentivize more effort, and what is the effect on risk-taking? We address these questions in the framework of our two-state model.

We begin by observing that a progressively more stringent threshold influences the desired allocation of effort and risk in two distinct ways. First, maintaining the existing level of compliance may require adjustments, potentially in both effort and risk-taking. Second, the higher threshold may prompt a shift in the chosen level of compliance itself, which in turn affects the preferred combination of effort and risk.

Proposition 2 shows that, as long as the compliance level remains unchanged, a more demanding threshold always induces higher effort. Its effect on risk-taking, however, is ambiguous and depends on the compliance regime: under partial compliance, a higher threshold increases risk-taking, whereas under full compliance it reduces it. This contrast arises from the different interaction between effort and risk-taking under partial versus full compliance.[4] Proposition 4 characterizes the optimal transitions across compliance regimes as the threshold increases. In particular, as the threshold becomes more demanding, agents optimally move from full compliance to partial compliance, and ultimately – when the threshold is very high – to no compliance. Finally, we compare the optimal levels of effort and risk-taking with their status quo counterparts, and we conclude the section with a fully worked out CRRA example that illustrates the potentially non-monotonic patterns that may arise as compliance shifts.

[3] A sufficient condition is that the returns to the risky asset have enough spread relative to the average return. In the case of a CARA function v, this condition reads $\frac{R}{r} < \frac{(1+R)^\rho}{(1-r)^\rho}$, which, if $\rho = 1$, reduces to the intuitive condition $\frac{R-r}{rR} < 2$.

[4] A caveat is in order. Throughout the discussion of the effects of a change in the threshold on effort and risk-taking, we implicitly focus on cases in which the threshold is binding (i.e., there is no slack). When the optimal response to the threshold involves slackness – either in the bad or in the good state – small changes in the threshold (that is, changes that do not alter the desired compliance regime) have no effect on either risk or effort. In such cases, the solution with slackness is interior and therefore, by definition, independent of the threshold constraint.

4.1. Effort and risk-taking with different thresholds

4.1.1. Comparative statics conditional on the compliance plan

The analysis of this section keeps the compliance level fixed. Since the optimal choices in regions Ω_1, Ω_2^o and Ω_3^o do not depend on the threshold level, we focus on regions $\bar{\Omega}_2$ and $\bar{\Omega}_3$, where the threshold binds either in the good or in the bad state, respectively. The next proposition shows that both effort and risk-taking respond monotonically to increases in the threshold level. However, while effort always increases, the effect on risk-taking depends on the compliance level.

Proposition 2. *The optimal choice follows the following pattern with respect to Y:*

1. *Under partial compliance (region $\bar{\Omega}_2$), both effort and risk-taking monotonically increase with the threshold level Y;*
2. *Under full compliance (region $\bar{\Omega}_3$), effort monotonically increases with the threshold level Y, while risk-taking monotonically reduces when Y increases.*

Point 1 indicates that, as the threshold becomes more demanding, partial compliance is optimally achieved through increases in both effort and risk-taking. This follows from the fact that effort and risk act as substitutes when the threshold is met only in the good state: higher risk-taking reduces the amount of effort needed. Therefore, as the threshold rises, some of the additional (and increasingly costly) effort required to comply in the good state is optimally replaced by greater risk-taking.

Point 2 tells us that under full compliance an increase in the threshold level causes a reduction in risk-taking and an increase in effort. This second result can be understood in terms of the "complements" property of effort and risk when the threshold is met in the bad state: higher risk-taking increases the required effort. The optimal reduction in risk-taking is therefore aimed at limiting the amount of effort which is needed to comply in the bad state. Interestingly, this result does not stem from risk aversion (in fact, in this plan the threshold is met for sure), but from a desire to equate the marginal cost of effort and the expected marginal utility from consumption. As long as the marginal cost of effort is increasing, the result holds for the case of a risk-neutral decision-maker as well.

4.1.2. Switches in the optimal plan

We start by defining maximizers for three different utility maximization problems, that will be useful for deriving the results that follow. The pair (α_1, e_1) maximizes expected utility under a certain loss; (α_2, e_2) maximizes expected utility when the agent faces a loss in the bad state and a gain in the good state; and (α_3, e_3) maximizes expected utility under a certain gain.

$$\begin{cases} (\alpha_1, e_1) = \text{argmax}[E_1\tilde{U}\,(\alpha, e)] \\ (\alpha_2, e_2) = \text{argmax}[E_2\tilde{U}\,(\alpha, e)] \\ (\alpha_3, e_3) = \text{argmax}[E_3\tilde{U}\,(\alpha, e)] \end{cases}$$

The portfolio realizations given these effort/risk combinations are denoted by: (y_L^1, y_H^1), (y_L^2, y_H^2) and (y_L^3, y_H^3), respectively. Note that these maximization problems are not conditional on a threshold level (although the solutions do depend on the magnitudes of G and L). The following proposition compares these solutions to the status quo levels of effort and risk-taking (under either CARA or DARA risk attitudes):

Proposition 3. *Let v be either DARA or CARA and $G = L$. Then:*

1. *$e_1 > e^*$ and $\alpha_1 < \alpha^*$;*
2. *$e_2 > e^*$ and $\alpha_2 < \alpha^*$;*
3. *$e_3 < e^*$ and $\alpha_3 > \alpha^*$.*

The first and last results follow from the fact that marginal utility and marginal cost must be equated. First, this implies that it is optimal to reduce effort when expecting a certain gain, and to increase effort when expecting a certain loss. The reduction in risk-taking under a certain loss (and the increase under a certain gain) is instead a direct consequence of the assumption of decreasing relative risk aversion. In the second result, both the decrease in risk-taking and the increase in effort are a direct consequence of the fact that both DARA and CARA admit precautionary savings,[5] and that while savings increase with variability, risk-taking decreases with variability (a detailed proof can be found in the Appendix).

In the next proposition, we record the ordering in the realizations induced by the above maximizers.

Lemma 1. *Let v be either CARA or DARA. Then:*

$$y_L^3 < y_L^* < y_L^2 < e^* < y_H^2 < y_H^* \tag{21}$$

We are now ready to characterize the switches in compliance levels that optimally occur as a result of an increase in the threshold level. The general intuition is that more demanding thresholds require larger distortions in effort and/or risk (with respect to the status quo), and thus weaken the incentives to comply.

[5] In the present analogy, effort plays the role of savings for an uncertain investment.

Proposition 4. *Let (α', e') denote the optimal solution of problem (15) with threshold Y.*

1. *If $Y < y_L^3$, then $(\alpha', e') \in \Omega_3^\circ$.*
2. *If $y_L^3 \leq Y < y_L^2$, then $(\alpha', e') \in \bar{\Omega}_3$.*
3. *If $y_L^2 \leq Y < y_H^2$, then $(\alpha', e') \in \{\Omega_2^\circ, \bar{\Omega}_3\}$.*
4. *If $y_H^2 \leq Y$, then $(\alpha', e') \in \{\bar{\Omega}_2, \bar{\Omega}_3, \Omega_1\}$.*

Moreover, as Y increases, the transition between choice regions must be consistent with the following ordering:

$$\Omega_3^\circ \to \bar{\Omega}_3 \to \Omega_2^\circ \to \bar{\Omega}_2 \to \Omega_1. \tag{22}$$

We see that when the threshold is larger than y_H^2, compliance is never "slack", in the sense that the threshold is met exactly (either in the bad or in the good state, depending on the level of compliance). A complete characterization of such switches depends on the specific form of the utility and cost functions. In Section 4.3 we fully characterize the compliance pattern (and the associated levels of effort and risk-taking) for different threshold levels in the context of a CRRA example.

4.2. Comparison with the status quo

The previous section looked at how the level of the threshold affects effort and risk-taking. We now examine the very impact of introducing a threshold—specifically, how effort and risk-taking compare to the "status quo" levels. In the following propositions, we explore both full and partial compliance (assuming that the threshold is met without slackness). These are the only relevant compliance scenarios when $Y > y_H^2$ (as shown in Proposition 4). Cases where the threshold is met with slackness are discussed in Proposition 3. We then draw conclusions about the effects of thresholds of varying magnitudes in the subsequent corollary.

Proposition 5. *Full compliance prescribes the following changes compared to the status quo:*

1. *More effort and less risk-taking for $Y > e^*$. For high thresholds ($Y > y_H^*$) risk-taking becomes zero.*
2. *Less effort and less risk-taking for low thresholds ($y_L^* < Y < e^*$).*
3. *Less effort and more risk-taking for very low thresholds ($Y \leq y_L^*$).*

The decrease in risk-taking for moderate to high thresholds does not arise from risk aversion. In fact, the result holds even under risk neutrality (i.e., with a linear v function).[6] Instead, the key mechanism behind point 1 is the complementarity between effort and risk under full compliance (see the discussion following Proposition 2). More precisely, full compliance prescribes limiting the necessary increase in effort by means of a reduction in risk-taking. This mechanism leads to a reduction in risk relative to the status quo, and to the complete elimination of risk for sufficiently high thresholds. It can, therefore, be argued that, when full compliance is the optimal response, moderate to high thresholds effectively function as a form of induced risk aversion.

Proposition 6. *Partial compliance always prescribes more effort than in the status quo. For high enough thresholds ($Y > y_H^* + \left(\frac{1-\alpha^* r}{1+\alpha^* R}\right) L$) risk-taking also increases compared to the status quo.*

The increase in risk-taking at high thresholds can again be traced back to the way effort and risk interact. Under partial compliance, effort and risk-taking are substitutes: relying more on one allows for less reliance on the other. When the threshold becomes very demanding, exerting additional effort becomes increasingly costly. At that point, it becomes optimal to reduce effort by taking on more risk instead. It is important to note that the level of Y at which risk-taking rises above the status quo becomes lower when the loss L is larger, or when the agent starts from a low level of risk-taking in the status quo (which reflects low risk aversion). In other words, both a larger potential loss and a less risk-averse baseline make the agent switch to higher risk-taking sooner.

The next corollary brings together the results obtained in Propositions 3 to 6 and Lemma 1 to assess the impact of the very introduction of thresholds of different magnitudes on effort and risk-taking.

Corollary 1. *Introducing a threshold has the following effects:*

1. *Effort is lower than in the status quo for low thresholds. Effort is instead higher for moderate and high threshold levels (that is, for $y_L^2 < Y < y_H^2$);*
2. *Risk-taking is higher than in the status quo for low thresholds ($Y < y_L^2$). At intermediate thresholds (up to the threshold level $Y = y_H^* + \left(\frac{1-\alpha^* r}{1+\alpha^* R}\right) L$) risk-taking drops below the status quo level. At higher threshold levels, risk-taking peaks above the status quo level, up to the point where there is no compliance. Beyond such thresholds risk-taking drops again below the status quo level.*

In the next section, we illustrate in the context of a CRRA example both the comparative statics within a given compliance region, and the shifts between compliance regions that arise as the threshold level increases.

[6] A risk-neutral agent would set $\bar{\alpha}_3 = 1$ in the status quo, and would still decrease risk to a level $\bar{\alpha}_3 < 1$ in the attempt to meet the threshold in both states.

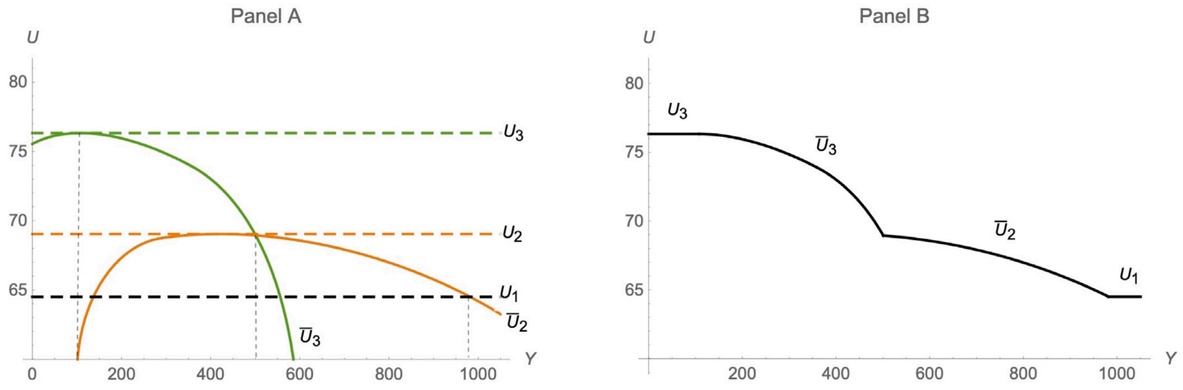

Fig. 1. Welfare as a function of the threshold level. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

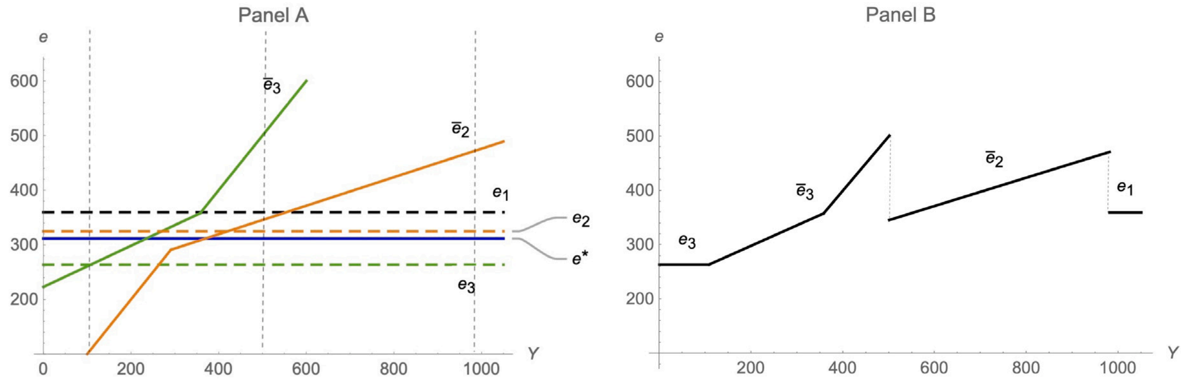

Fig. 2. Effort as a function of the threshold level. Panel A: optimal effort by compliance region. Panel B: optimal realized effort.

4.3. A CRRA example

Set $v(y) = 2y^{\frac{1}{2}}$ and $C(e) = A - 2(600 - e)^{\frac{1}{2}}$, with fixed cost $A > 2(600)^{\frac{1}{2}}$. The cost function is positive, strictly convex and strictly increasing in effort e. We set $r = 0.8$, $R = 1.4$, $L = G = 100$. We shall denote by $(\bar{e}_2, \bar{\alpha}_2)$ and $(\bar{e}_3, \bar{\alpha}_3)$ the optimal solutions in $\bar{\Omega}_2$ and $\bar{\Omega}_3$, respectively. In Fig. 1 we report the utility levels attained by means of the maximizers (α_1, e_1), (α_2, e_2), $(\bar{\alpha}_2, \bar{e}_2)$, (α_3, e_3) and $(\bar{\alpha}_3, \bar{e}_3)$. Consistent with Proposition 4 (points 1. and 2.), the optimal compliance regime in the region $Y < 108 = y_L^3$ is full compliance (with slackness, i.e. Ω_3^0), while in the region $108 \leq Y < 270 = y_L^2$ it is optimal to fully comply with no slackness (i.e. $\bar{\Omega}_3$).

In the region $y_L^2 = 270 \leq Y < 420 = y_H^2$ the effort/risk combination (α_2, e_2) is never an optimal strategy, since the partial compliance combination $(\bar{\alpha}_3, \bar{e}_3)$ dominates in this region in terms of expected utility (see Proposition 4 point 3). Full compliance remains optimal until the threshold level $Y = 501$, after which partial compliance becomes optimal. Finally, for thresholds $Y > 981$ it is optimal to give up the threshold altogether. The switches in the optimal compliance level are marked by the vertical dotted lines in Panel A of Fig. 1; the curve in Panel B highlights the realized (maximal) welfare as the threshold level increases.

Keeping the switches in compliance in mind, let us now look at the patterns of effort and risk-taking as the threshold level increases.

Fig. 2 displays the optimal effort levels as the threshold increases. As in Fig. 1, the vertical lines in Panel A mark the switches in the optimal compliance regions. The dashed horizontal lines refer to the regions of compliance with slackness. In Panel B we record the optimal level of effort only in the optimal compliance regions. We see that although the optimal effort is weakly increasing for any given fixed compliance region (see Proposition 2), switches in compliance regions are responsible for a non-monotonic pattern. In particular, when the optimal plan switches from full to partial compliance (at $Y = 501$), effort drops suddenly as a consequence of the switch in the target from the low to the high state realization. Effort then starts to increase again until a new switch in the optimal plan occurs at $Y = 981$, after which the threshold is given up altogether and effort drops to the optimal level in anticipation of a certain loss. Note that this level of effort (red line) is above the "status quo" (blue line), consistent with Proposition 3.

Fig. 3 displays optimal risk-taking as a function of the threshold. We observe a decreasing pattern under full compliance (see Proposition 2). At $Y = 501$, the switch from full to partial compliance is responsible for the abrupt increase in risk-taking. Risk-taking then continues to increase up to the point where giving up the threshold becomes optimal ($Y = 981$), at which point risk-taking drops due to the anticipation of a certain loss. This lower level of risk-taking is below the "status quo" level, consistent with Proposition 3.

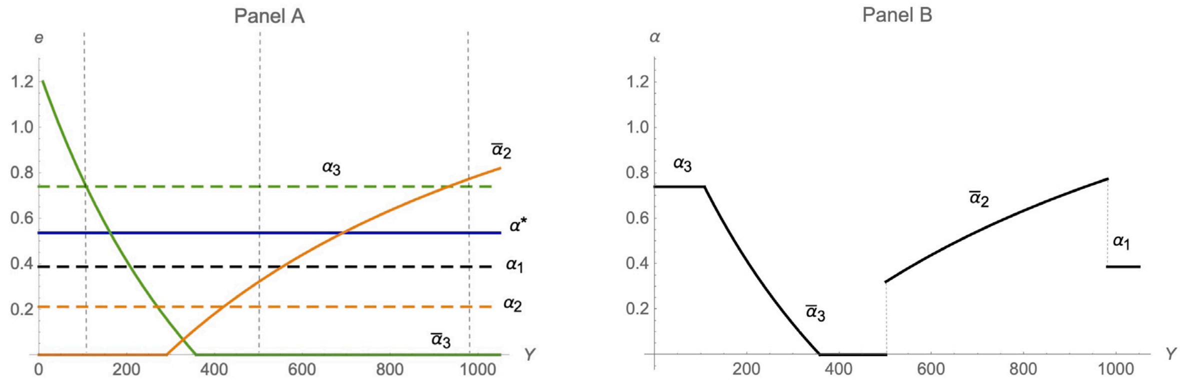

Fig. 3. Risk-taking as a function of the threshold level. Bottom panel: black line shows effort under optimal plans.

5. Experimental evidence

This section studies the main insights of the theory above using an experiment with 1092 participants. In the experiment, participants decide how much effort to exert and whether to take risk, both in situations without a threshold and in situations where they receive an additional payment for reaching a specified threshold. To make the task easier to understand, we simplify the theoretical setting in two respects. First, we set $L = 0$, so that participants do not face any loss from failing to reach the threshold. Second, instead of choosing a risk level from a continuous scale, participants make a binary choice of whether or not to take risk. These simplifications make the experimental design more transparent and easier to understand, while preserving the core mechanisms of the model.

Setting $L = 0$ implies that a few aspects of the theoretical framework cannot be examined directly in the experiment. In particular, the design does not allow us to study behavior in situations where a threshold is exceeded with slackness (only relevant for low thresholds). Importantly, however, all predictions concerning full compliance and partial compliance with a binding threshold remain directly applicable and testable. Modeling risk-taking as a binary decision means that in any partial compliance scenario risk-taking is necessarily maximal. Even in this simplified setting, the design continues to allow for a clean test of the central theoretical prediction regarding substitution between effort and risk. The trade-off is now mainly reflected in the transition from full to partial compliance, providing a simpler but transparent channel through which the substitution can be identified.

5.1. The experiment

Design. The experiment consisted of 20 rounds. In each round, we asked participants how many "tasks" (between 0 and 100) they would like to solve for a given payment scheme. Then, the computer randomly picked one of the 20 rounds and participants had to solve the number of tasks that they selected in that round for the payment scheme offered in that round.

Table 1 shows the 20 choices that participants made throughout the experiment. Below, we describe each of them:

- In Round 1 (R1), participants had to choose how many tasks to solve at the price of GBP 0.1. This choice helps us to measure the elasticity of substitution between effort and money for each subject. As such, we will use it to classify participants based on their preferences.
- In Round 2 to 5 (R2–R5), participants faced the same choice as in R1 and, in addition, were paid GBP 3 if their payment reached a payoff of GBP 2, 4, 8, or 10 (which we call thresholds 20, 40, 80, and 100 respectively). For example, a participant who faced a threshold of 40 would be paid GBP 3.5 if she decided to solve 35 tasks, but she would be paid GBP 7.5 if she decided to solve 45 tasks (GBP 0.1 * 45 + 3). These rounds were randomized, such that participants faced the four different thresholds in random order.
- In Round 6 to 8 (R6–R8), participants made two choices, which were both displayed on the same screen. First, they decided how many tasks they wanted to solve. Second, they decided whether to take risk or not. If they did not take risk, they were paid GBP 0.1 as in R1. If, however, they took risk, they were paid GBP 0.2 with 50% chance and GBP X otherwise, where $X \in \{0; 0.025; 0.05\}$ across rounds. These rounds were randomized, such that participants faced the three different payment schemes in random order.
- In Round 9 to 20 (R9–R20), participants faced decisions in which there were both thresholds and the option to take risk. For each of the four thresholds in R2–R5 (GBP 2, 4, 8, and 10) and payment schemes in R6–R8 (where the bad state pays GBP 0.05, 0.025 or 0), participants chose how many tasks to solve and whether or not to take risk. These rounds were also randomized.

Using this structure offers two main advantages. First, we believe that incrementally adding elements to participants' decisions (threshold, risk, and their combination) is important for a clear understanding of the decision problem and for reducing confusion.

Table 1
Experimental design.

Round	Certain payment	Payment in good state	Payment in bad state	Threshold
R1	0.1	–	–	–
R2	0.1	–	–	20
R3	0.1	–	–	40
R4	0.1	–	–	80
R5	0.1	–	–	100
R6	0.1	0.2	0.05	–
R7	0.1	0.2	0.025	–
R8	0.1	0.2	0	–
R9	0.1	0.2	0.05	20
R10	0.1	0.2	0.025	20
R11	0.1	0.2	0	20
R12	0.1	0.2	0.05	40
R13	0.1	0.2	0.025	40
R14	0.1	0.2	0	40
R15	0.1	0.2	0.05	80
R16	0.1	0.2	0.025	80
R17	0.1	0.2	0	80
R18	0.1	0.2	0.05	100
R19	0.1	0.2	0.025	100
R20	0.1	0.2	0	100

Note: Description of the payment schemes in the 20 rounds. In R1, participants chose how many tasks to solve for a payment of GBP 0.1 per task. In R2–R5, participants were additionally paid GBP 3 for reaching a threshold. In R6–R8, participants chose (1) how many tasks to solve and (2) whether to be paid GBP 0.1 per task or with a lottery yielding GBP 0.2 per task with 50% chance and GBP 0/0.025/0.05 otherwise. In R9–R20, participants chose again whether to be paid GBP 0.1 per task with certainty or with the lottery, and additionally receive GBP 3 for reaching a threshold. The order in which participants viewed R2–R5 was randomized, as well as the order of R6–R8 and R9–R20.

Second, it enables us to exploit within-subject dimensions, which are crucial for identifying the heterogeneities predicted by our theory. A potential issue with this design could arise if participants systematically chose differently across their decisions. However, when analyzing the randomized R2–R5, R6–R8, and R9–R20 decisions, we do not find evidence of order effects in our experiment.[7]

The task. The task involved correctly counting the number of zeroes in 10 × 10 matrices (Abeler et al., 2011). For each matrix, participants were asked to enter the number of zeroes in the matrix. If they entered the correct number, the task would count as solved and participants moved to the next one. If they did not enter the correct number, participants had two additional tries. If they failed the three attempts, they moved to the following matrix and the task did not count as solved.

Counting zeroes is a tedious and repetitive task that typically takes about 25 s. Before participants played their 20 rounds, they all had to correctly solve three practice tasks to understand how costly it is for them to solve these tasks.

Procedures. We fielded the experiment in April 2021 on the platform Prolific. Participants were told that they would be paid GBP 2 to answer a set of questions that would take about 10–15 min. We told them that, afterwards, they would have the possibility to keep solving tasks (for up to one hour) for additional payments.

Once participants had read the instructions and solved the three practice tasks, they played the 20 rounds. Before R1, R2–R5, R6–R8, and R9–R20, participants read instructions describing the upcoming rounds. To make sure that they understood them, they had to answer two control questions for each part of the experiment. They could only move forward when they answered these questions correctly.

At the end of the survey, we asked participants several background questions, including their sex, age, ethnicity, social class, education, math ability, ambition, and willingness to take risks.

The sample. We restricted the sample to UK residents only, but made no other sample restrictions. In the pre-registration, we promised a sample size of "At least 500 participants and up to the number that ensures 200 participants with R1 choices between 40 and 80 tasks".[8] In the end, we stopped data collection after we had 1092 participants, since we then reached 201 participants whose R1 choices were between 40 and 80. We will call the sample of 1092 participants the *full sample* and the sample of 201 participants the *pre-registered sample*. The pre-registration can be found in https://www.socialscienceregistry.org/trials/7536. Our special interest in the pre-registered sample stems from the fact that, for these participants, we can reasonably interpret the thresholds of 20 and

[7] For each of the decision sequences (R2–R5, R6–R8, and R9–R20), we tested whether the order in which participants completed the rounds affected their effort or risk choices. We found no systematic order effects, and all estimates from the main pre-registered tests remain virtually unchanged when controlling for order fixed effects.

[8] The sample size was determined based on a power analysis that used pilot data from 34 participants as seed data for the simulations (following the procedures in Campos-Mercade, 2024).

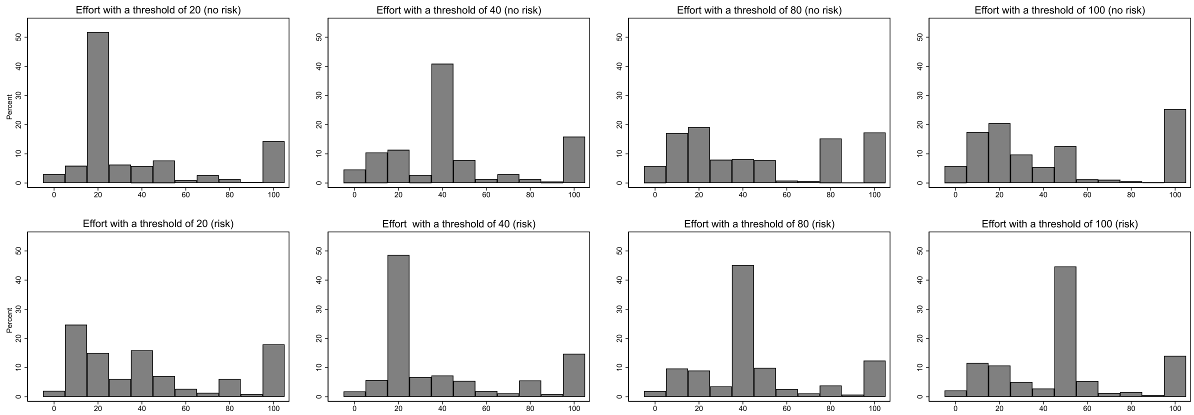

Fig. 4. Average effort choice by threshold level and by whether risk is taken (bottom panel) or not (top panel). Full sample.

40 as "low thresholds", and the thresholds of 80 and 100 as "high thresholds". Hence, this sample allows us to more clearly study the predictions of the model.[9]

Empirical questions. To find empirical support for the key theoretical insights and behavioral implications, we focus on the following three main empirical questions:

Q1: *Relevance of the threshold (pre-registered research question)*. Does the introduction of a threshold affect agents' behavior in terms of effort and risk-taking? Does the effect depend on the level of the threshold? Behavior without a threshold refers to R6–R8. Behavior with a threshold refers to R9–R20. We compare average effort levels and the share of participants taking risk both with and without threshold and across threshold levels (high vs. low thresholds).

Q2: *Heterogeneous effects for heterogeneous types.* Does the pattern by which the threshold affect effort and risk-taking vary across heterogeneous participants?

Q3: *Substitution vs. Complementarity*. Can we identify the key mechanism behind the comparative statics results of Section 4.1.1, that is, the different interplay between effort and risk-taking under different plans?

5.2. Summary statistics

Table 2 presents the summary statistics for all participants (column 1) and for the participants in the pre-registered sample (column 2). Specifically, it shows the main characteristics of the participants (age, gender, ethnicity, self-reported social class, and education) and also their average answers on R1 (basic piece rate) and R9–R20 (involving both thresholds and risk choice). Appendix Table B.1 includes summary statistics for the rest of the rounds.

Our sample is representative in terms of ethnicity, but contains more women and is younger and more educated compared to the UK population on average. The large sample size allows us to split the sample by these characteristics and we do not find significant differences in behavior across sub-samples. In particular, none of the treatment effects that we study is significantly different across gender and age. As for the pre-registered sample, there are somewhat fewer women compared to the full sample (t-test, $p = 0.080$), but otherwise it does not significantly differ from our main sample in terms of these demographic characteristics.

We find substantial heterogeneity in terms of both effort and risk choices across all the rounds. By definition, effort choice in R1 differs across the two samples and is substantially higher in the pre-registered sample with a mean of about 51 compared to a mean of about 37 in the entire sample. The table also shows that – on average – participants exert more effort with thresholds (R9–R20) than without thresholds (R1), especially for the full sample. Furthermore, the average amount of risk-taking differs substantially across rounds ranging from less than 20% of participants taking risk in R11 to more than 60% in R18.

To get a picture of participants' effort decisions, Fig. 4 shows the distribution for the full sample of participants' effort choices in rounds 9 to 20 by threshold level and by the decision of whether to take risk (bottom panel) or not (top panel). For each threshold, we pool decisions across the three risk returns. The figure shows that there is a considerable amount of heterogeneity with effort choices covering the entire range between 0 and 100.

It also shows the pattern that the decisions for participants who do not take risks spike at exactly the effort required to attain the threshold (20, 40, 80, and 100). This is most clearly visible for the lower thresholds of 20 and 40, but also for the high thresholds

[9] The reason we did not pre-register the full sample as our main sample of interest is that due to heterogeneity of preferences it is not possible to make general predictions. For example, a threshold of 40 may be very low for a participant who decides to solve 100 tasks in R1, while it may be very high for a participant who solves 10 tasks in R1. Hence, the model's predictions for one participant may be completely in conflict with those for a different participant.

Table 2
Summary statistics.

	(1)		(2)	
	Full sample		Pre-registered sample	
	Mean	Std. deviation	Mean	Std. deviation
Age	35.76	12.32	34.76	12.07
Female	.68	.47	.63	.48
White	.87	.33	.90	.31
Working class	.50	.50	.57	.50
University studies	.64	.48	.63	.48
Effort R1	37.04	32.57	50.85	8.63
Effort R9	41.40	30.97	49.05	24.88
Effort R10	40.47	32.28	49.55	27.77
Effort R11	36.89	30.77	43.05	24.89
Effort R12	45.31	31.40	55.16	24.23
Effort R13	41.44	28.98	48.89	20.17
Effort R14	41.26	29.04	48.18	20.85
Effort R15	45.39	31.00	56.34	22.08
Effort R16	45.32	31.48	57.65	22.19
Effort R17	45.42	32.18	58.59	22.73
Effort R18	46.55	32.26	57.05	21.05
Effort R19	45.96	32.89	57.47	23.03
Effort R20	45.24	33.52	58.00	23.38
Risk R9	.55	.50	.57	.50
Risk R10	.39	.49	.37	.48
Risk R11	.27	.44	.20	.40
Risk R12	.55	.50	.48	.50
Risk R13	.40	.49	.30	.46
Risk R14	.29	.46	.23	.42
Risk R15	.54	.50	.58	.49
Risk R16	.37	.48	.41	.49
Risk R17	.27	.44	.28	.45
Risk R18	.56	.50	.63	.48
Risk R19	.38	.48	.48	.50
Risk R20	.25	.44	.31	.46
N		1092		201

Note: R1 pays the basic piece-rate. Rounds 9–11 have a threshold of 20, R12–R14 of 40, R15–R17 of 80 and R18–R20 of 100. Within each of these the bad state pays GBP 0.05, 0.025 or 0, respectively.

of 80 and 100 we do see substantial mass points at exactly the threshold. Given the average effort choice of 37 in the absence of thresholds (the "status quo choice"), these higher thresholds require substantial adjustments by most participants.

By contrast, when participants do take risk, their decisions tend to spike at the level that allows them to just attain the threshold in the good state. This is 10 for the threshold of 20, 20 for the threshold of 40, 40 for the threshold of 80, and 50 for the threshold of 100. The bottom panel of Fig. 4 shows that decisions peak at these levels. Consistent with our theoretical prediction, risk-taking characterizes the choice of meeting the threshold in the good state only, while it is absent (or reduced) when meeting the threshold for sure.

5.3. Results

First, we perform our pre-registered analysis by restricting the sample to those participants who chose to solve between 40 and 80 tasks in Round 1 (Question 1). Second, we use all participants' answers in Round 1 to divide them into four types and explore whether their behavior is in line with the model (Question 2). Finally, we use all participants' answers to study the interplay between effort and risk-taking (Question 3).

We structure our results along the three testable empirical questions discussed above.

5.3.1. Relevance of the threshold for effort and risk-taking

To test the very effect of *introducing* a threshold, we compare participants' average decisions in R6–R8 (status quo treatments) with those in R9–R20 (threshold treatments). Across all rounds, we find that the introduction of thresholds does not significantly increase effort (mean difference = −0.70 tasks; $t(2411) = -1.56$, $p = 0.119$), but it significantly increases risk-taking by 4.15 percentage points ($t(2411) = 3.86$, $p < 0.001$).

These average effects are not very informative, however, since they do not take into account the role played by the threshold level. In our theory section, we have shown that the threshold level is supposed to affect effort and risk-taking in very different ways. We find that introducing a low threshold *decreases* effort by 4.97 tasks ($t(1205) = -7.79$, $p < 0.001$) but does not affect risk-taking ($t(1205) = -0.22$, $p = 0.83$). However, introducing a high threshold *increases* effort by 3.57 tasks ($t(1205) = 5.88$, $p < 0.001$) and

Table 3
Effect of a high threshold on effort and risk-taking.

	Effort			Risk-taking		
	(1)	(2)	(3)	(4)	(5)	(6)
High threshold	8.536***	8.536***	8.753***	0.090***	0.090***	0.098***
	(0.878)	(0.881)	(0.916)	(0.022)	(0.022)	(0.023)
Observations	2412	2412	2412	2412	2412	2412
Mean outcome	53.249	53.249	53.249	0.405	0.405	0.405
Pre-registered controls		X	X	X	X	X
Risk-taking control			X			
Effort control						X

Note: Main pre-registered test in which we compare the difference on effort (columns 1 and 2) and risk-taking (columns 4 and 5) of a high threshold with respect to a low threshold. The estimates correspond to an OLS regression with standard errors clustered at the participant level. Columns 1 and 2 do not have any controls, and Columns 4 and 5 control for gender, age, social class, ethnicity, and education. Column 3 conducts the pre-registered analysis for effort while controlling for risk-taking, and Column 6 conducts the pre-registered analysis for risk-taking while controlling for effort. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

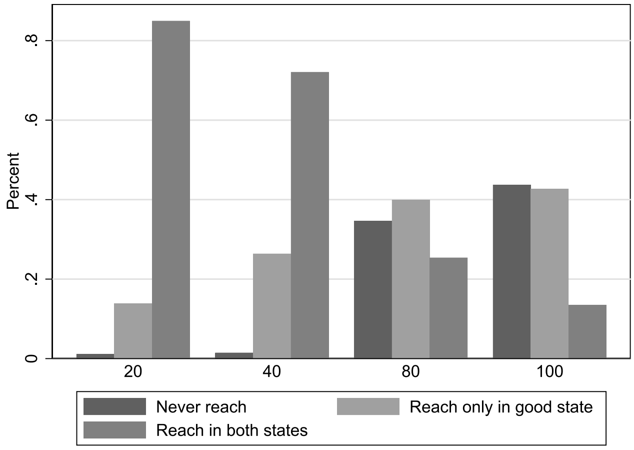

Fig. 5. Percentage of participants who never reach the threshold, reach it in the good state only, or reach it in both states, depending on the level of the threshold.

risk-taking by 8.62 percentage points ($t(1205) = 5.76$, $p < 0.001$). This evidence is in line with our results in Section 4.2, where very low thresholds are shown to reduce effort (compared to the status quo) through an income effect, while high thresholds increase both effort and risk-taking due to a substitution effect. This evidence also suggests that behaviors react to the magnitude of the threshold in line with the comparative statics of Section 4.2.

To investigate this issue in more depth, we need to account for the fact that the same threshold of 80 could be perceived as "high" for someone who would choose 50 tasks in the status quo, but not for someone whose status quo choice is 100 tasks. For this reason, we pre-registered that we would use participants' decisions in R1 to divide them into similar types. In particular, we pre-registered that we would use participants who decided to solve between 40 and 80 tasks in R1. For these participants, the thresholds of 20 and 40 are lower than what they would normally choose, while the thresholds of 80 and 100 are higher.

Table 3 shows that the threshold's level has a significant effect on both effort and risk-taking. On average, participants solved 8.5 more tasks ($p < 0.001$) and are 9 percentage points more likely to take risk ($p < 0.001$) when they face high thresholds. Higher thresholds thus increase both effort and risk-taking compared to lower thresholds, consistent with Proposition 2. The results are similar when we control for risk-taking (when explaining effort) and for effort (when explaining risk-taking).

Finally, we test whether changes in the decision to meet the threshold are in line with the intuitive prediction of Proposition 4. We record the optimal plans in Fig. 5 as a function of the threshold level. Specifically, we record the percentage of participants who meet the threshold in both states of the world, in only the good state, or never.

Consistent with Proposition 4, low thresholds are typically reached in both states of the world, while the higher the threshold the more likely it is that it is only reached in the good state, and eventually that it is never met. Note that all these three plans are feasible for participants, who are always able to meet all threshold levels in both states of the world (although they may not choose to do so). Fig. 5 also suggests that participants decide to reach high thresholds not via a (potentially large) increase in effort; instead, they choose to take risk and decide to reach the threshold in the good state only.

We can summarize these findings in the following:

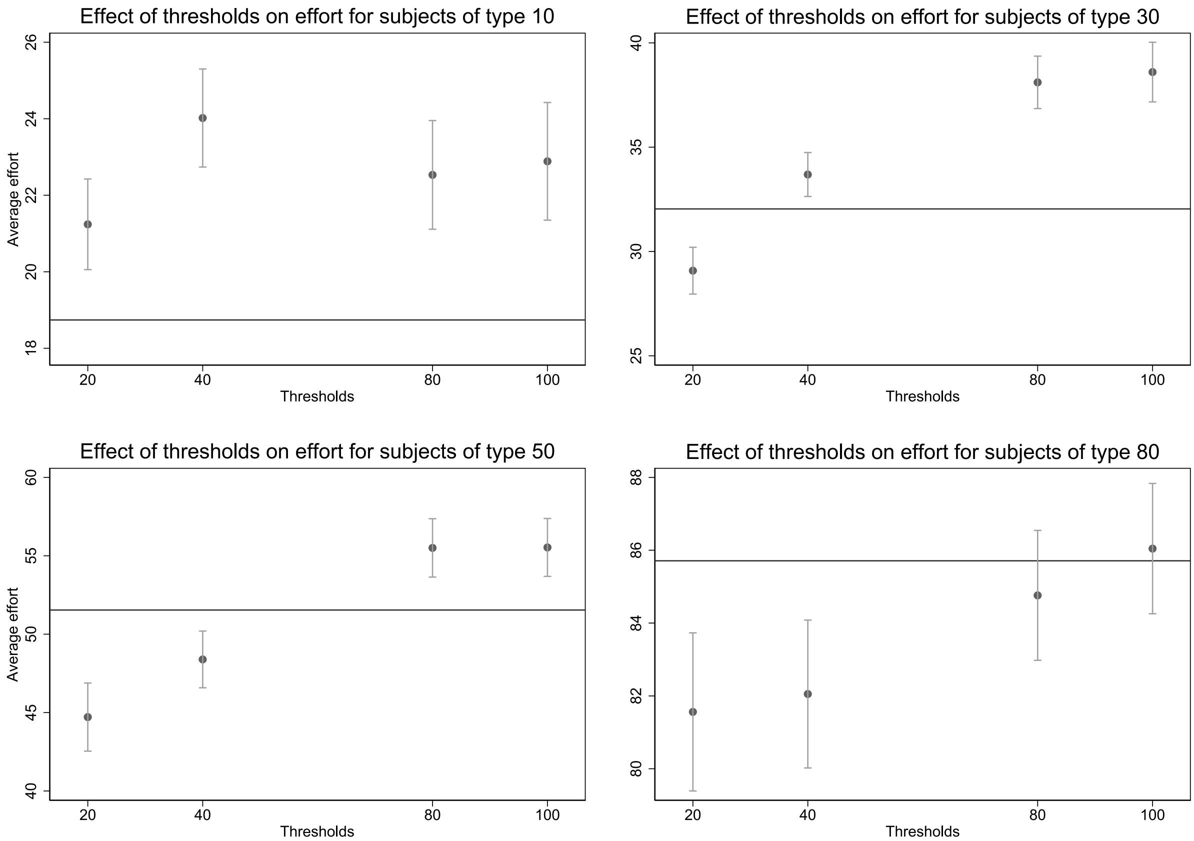

Fig. 6. Effort levels by type of participant and threshold level.

Result 1. *The introduction of a threshold affects effort and risk choice and the direction of the effect depends on the threshold level. As the threshold level increases (i) participants who opt for partial compliance tend to exert more effort, and a higher proportion of them take risk; (ii) participants who opt for full compliance tend to exert more effort, and fewer of them take risk. Moreover, compared to the case of no threshold, introducing high thresholds increases both effort and risk-taking.*

5.3.2. Heterogeneous effects for heterogeneous types

The model predicts that the effect of a threshold depends on how demanding it is relative to a person's "status quo" level of performance. For participants who normally exert little effort, even a modest threshold may feel high and motivate substantial additional effort; for those who already work hard, the same threshold may appear low and provide little or even negative incentive. To explore this heterogeneity, we divide participants into four types according to their effort choice in the baseline round R1 (without thresholds or risk). More concretely, we divide the sample between those whose R1 decision is between 0 and 19 (*Type 10*), between 20 and 39 (*Type 30*), between 40 and 59 (*Type 50*), and more than 60 (*Type 80*).[10] This classification offers a simple proxy for participants' perceived distance to each threshold.

Fig. 6 shows the average effort by type across thresholds of 20, 40, 80 and 100. For comparison, the horizontal line represents the average effort level that participants exerted in the threshold-free rounds (R6–R8).

For low-effort participants (Type 10), the introduction of a threshold strongly increases effort: even the lowest threshold (20) has a strong positive effect, and the 40 threshold raises it further. However, once thresholds become very demanding (80 or 100), average effort stops rising and even declines slightly—consistent with the idea that most participants give up on pursuing these higher thresholds.

For intermediate types (Type 30 and Type 50), effort rises when thresholds move from 20 to 80 but plateaus at 100. For the highest-effort group (Type 80), thresholds of 20 or 40 actually reduce effort relative to the baseline, and only very high thresholds

[10] As secondary tests, we pre-registered that we would use a much more complex classification for threshold difficulty. That analysis (not reported) leads to very similar intuitions as those that we report in this section. The advantage of using this simpler analysis is that the classification based on R1 is closer to the one we use for the pre-registered analysis and much more straightforward. Additionally, using this classification helps us compare effect sizes of the thresholds as they involve comparisons within the same sample.

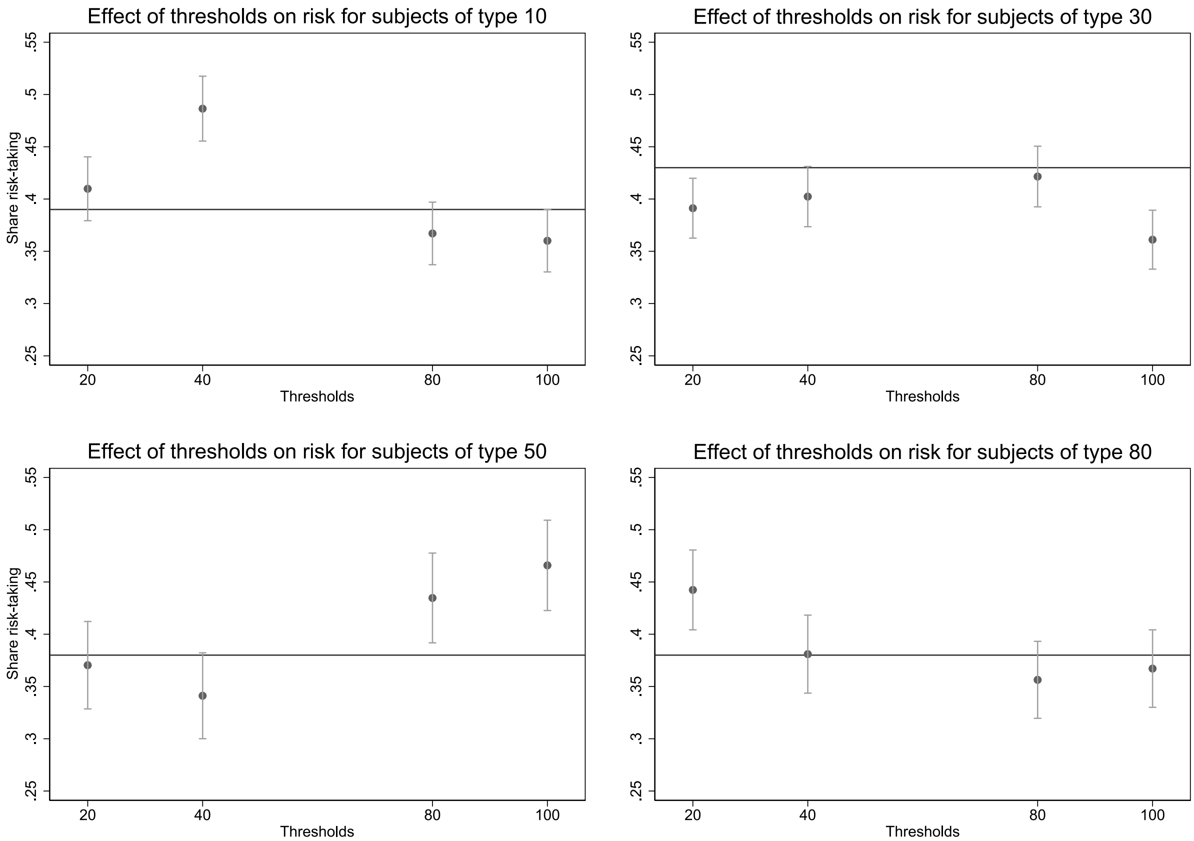

Fig. 7. Share of risk choices by type of participant and threshold level.

(80 or 100) restore effort to previous levels. This pattern shows how thresholds that are too low relative to a person's status quo can reduce effort. In short, thresholds increase effort among low-effort participants but may lower effort among high-effort participants, implying that optimal threshold design must consider where individuals start from.

Turning to risk-taking (Fig. 7), we again find heterogeneous responses. Participants whose baseline effort lies somewhat below a threshold – the group for whom the threshold is challenging but attainable – show the largest increase in risk-taking. By contrast, participants whose baseline effort already exceeds the threshold show little change or even small reductions in risk-taking. Overall, in line with the intuition developed in the model, thresholds tend to raise risk-taking for those just below the target and have negligible or negative effects otherwise.

Taken together, these heterogeneous responses help interpret the average effects seen earlier. When aggregated, thresholds seem to modestly increase both effort and risk-taking, but this masks substantial variation: thresholds generally motivate low-effort individuals while reducing effort among high-effort individuals. This pattern underscores that thresholds are not uniformly motivating instruments—their effectiveness depends on whom they target and how far the target lies from individuals' status quo behavior.

Overall, these heterogeneous patterns are in line with the model's predictions. Thresholds raise effort mainly among individuals who exert little effort in the status quo, for whom meeting the threshold requires a clear adjustment in behavior. Among those already working hard, by contrast, thresholds are easily met and therefore elicit little change or even a slight reduction in effort. For risk-taking, increases are most visible among participants whose baseline effort lies just below the threshold – those for whom taking risk can help reach it – while for others the effects on risk are small or even negative.

Result 2. *Thresholds increase effort among participants who exert little effort in the status quo but have smaller or even negative effects among those who already exert high effort. Risk-taking rises mainly for participants whose baseline effort is just below the threshold, while for others the effect is close to zero or slightly negative.*

5.3.3. Substitution vs. complementarity in effort and risk-taking

We finally address our third research question, concerning the underlying mechanism that triggers the increase in risk-taking (and the parallel slowdown in the extra investment in effort) as the threshold becomes more demanding. As discussed after Proposition 2, the switch from full compliance (region $\bar{\Omega}_3$) to partial compliance (region $\bar{\Omega}_2$) changes the interplay between effort and risk in

Table 4
Effort and risk by threshold level and optimal region.

		Threshold level			
		20	40	80	100
No compliance	Mean effort	6.69	15.34	24.46	27.74
	% Risk	11%	17%	27%	32%
	% Δ Risk	−13%	−43%	−21%	−12%
	ρ	-0.138^{***}	-0.216^{***}	-0.192^{***}	-0.115^{***}
	N	2114	6078	10 534	11 985
Partial compliance	Mean effort	31.96	44.27	63.36	73.69
	% Risk	100%	100%	100%	100%
	% Δ Risk	+40%	+35%	+32%	+31%
	ρ	0.004	-0.047^{***}	-0.092^{***}	-0.108^{***}
	N	3213	3885	3036	2102
Full compliance	Mean effort	54.26	72.12	96.43	100
	% Risk	19%	9%	0%	0%
	% Δ Risk	−13%	−18%	−19%	−18%
	ρ	0.146^{***}	0.182^{***}	0.032	–
	N	9816	5351	3511	1787

Note: Mean Effort, % of participants who choose to take risk (% Risk), percent increase in risk-takers compared to the equivalent case without thresholds (% Δ Risk), pairwise correlation between effort and Δ Risk and number of observations (N) separately for different levels of the threshold and conditional on whether threshold is never met, met in the good state only or met in both states. This table uses the Full sample. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

attaining a given threshold, from complements (more risk requires more effort), to substitutes (more risk requires less effort). Since effort is increasingly costly, optimality requires limiting the extra investment in effort, and this is accomplished by means of opposite changes in risk-taking, depending on the plan: a decrease in $\bar{\Omega}_3$, an increase in $\bar{\Omega}_2$.

Table 4 provides evidence of this shift in the interplay between effort and risk-taking. It summarizes a number of decisions depending on the threshold level and conditional on the plan region that participants have adopted concerning the threshold: no compliance, partial compliance, or full compliance.

The table shows that mean effort increases as the threshold level increases, for each of the three compliance plans. Risk-taking increases with respect to the status quo (treatments without a threshold) under partial compliance, and decreases with respect to the status quo under full compliance. Also, risk-taking decreases with the threshold level under full compliance (with a null figure for risk-taking at threshold levels 80 and 100). All this is consistent with our results from Sections 4.1.1 and 4.2.

To get direct evidence of the different interplays between effort and risk-taking, we also look at the correlation coefficient ρ between effort and risk-taking. The figures show that under partial compliance the increase in risk-taking comes with a decrease in effort ($\rho < 0$ for all threshold levels). This is in line with the fact that effort and risk-taking move in opposite directions. By contrast, under full compliance $\rho > 0$, consistent with the fact that effort and risk-taking move in the same direction.

Another way of assessing the role of substitution between effort and risk-taking under partial compliance is to look at the effect of thresholds when there is no option to increase risk. In Appendix Table B.1 we show mean effort in R2–R5 where a threshold of 20, 40, 80 or 100 is introduced, but where there is no option to take risk. Introducing a threshold of 40 or higher increases effort substantially compared to the baseline without threshold, with the highest level of effort reached for thresholds of 80 and 100, for which effort levels do not differ significantly. Interestingly, in these two rounds effort levels are higher also when compared to any of the rounds (R6–R20) where risk-taking is possible. This is further evidence that when agents do have the possibility to take risks, some decide to meet the threshold only in the good state, causing a less pronounced increase in effort through the substitution effect.

Finally, we find similar results when restricting the analysis to the pre-registered sample instead of the whole sample as done in Table 4. On average, for participants who reach the threshold in both states, the correlation between effort and risk-taking is 0.192 ($p < 0.001$). Hence, risk and effort are complements in this region. However, if we consider all the participants who would reach the threshold in the good state if they took the risk, the correlation is −0.115 ($p < 0.001$), indicating that risk and effort are substitutes in this region.[11]

Result 3. *We find evidence that the interplay of risk and effort changes with the optimal plan. While they are negatively correlated under partial compliance, they are positively correlated under full compliance.*

[11] As a tertiary test, we pre-registered that we would study whether participants' decisions are different based on their background variables, and in particular gender. We find very little differences across all background variables that we measure. While women are less willing to exert effort in the baseline, with 69% of women classified as either type 10 or type 30 as opposed to only 55% of men, conditional on type, thresholds have a similar effect for both men and women. Similarly, we find no systematic gender differences on risk-taking: In an OLS regression where we explain risk-taking with the type of threshold (threshold 20, 40, 80, or 100), the type of subject (type 10, 30, 50, or 80), and a gender dummy, we find that females are 1.2% more likely to take on risk, but this is not statistically significantly different from zero. Hence, in contrast to previous work (e.g., Charness and Gneezy 2012), we do not find evidence that women are more risk-averse in our setting.

As in the pre-registered sample, we can also use the full sample to test one of the key implications of the model: effort and risk are complements when thresholds are met with certainty, and substitutes when thresholds are met in the good state. Once again, we find support for this hypothesis. When participants reach the threshold with certainty, the correlation between effort and risk is 0.182 ($p < 0.001$). However, when participants reach the threshold only in the good state, the correlation is −0.039 ($p < 0.001$).

6. Concluding remarks

Performance thresholds are typically introduced as contractual devices to incentivize effort, but they can also affect risk-taking. In a theoretical model and a controlled online experiment, we show that thresholds influence both effort and risk, and that their level is crucial for determining the direction and magnitude of these effects. More precisely, the distance between the threshold and the status quo levels of performance is the key driver of an agent's behavior, both in terms of compliance with the threshold and in terms of effort and risk-taking decisions.

What policy implications can we draw from our analysis? One key lesson from our analysis is that the effect of a threshold dramatically depends on the "type" of agent—that is, on whether the threshold is perceived by the agent as low, intermediate, or high. An economic plan based on the imposition of a threshold should therefore account for two features of optimal behavior. First, different agents may react differently to the same threshold, both in terms of effort and risk-taking, simply because they would optimally take different decisions regarding whether and how to meet the threshold. Second, an increase in the threshold level may have the "unintended" effect of decreasing effort and increasing risk-taking. Being aware of these behavioral features is key both to the targeting of the set of agents subject to the threshold, and to the fine-tuning of the threshold level.

Although our analysis lends itself to a variety of economic, social, and political applications – including health, education, finance, and international cooperation – some applications may require extensions of the basic framework. In economic unions, for instance, both the gains from membership and the losses from exclusion typically depend on how many, and which, countries join the union. In terms of the present analysis, this would imply that the terms G and L depend on the number of agents who meet the threshold. Such a modification would introduce a strategic element into the model, affecting both effort and risk-taking in equilibrium. Applications in environmental economics would also suggest that the threshold level itself may be uncertain (consider, for instance, critical temperature increases that trigger severe climate consequences). While we abstract from these issues in the present paper in order to focus on the basic interplay between risk-taking and effort, these extensions represent promising avenues for future research.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.euroecorev.2026.105363.

Data availability

Data will be made available on request.

References

Abeler, J., Falk, A., Goette, L., Huffman, D., 2011. Reference points and effort provision. Am. Econ. Rev. 101 (2), 470–492.
Amihud, Y., Lev, B., 1981. Risk reduction as a managerial motive for conglomerate mergers. Bell J. Econ..
Andersson, O., Holm, H.J., Wengström, E., 2020. Grind or gamble? An experiment on effort and spread seeking in contests. Econ. Inq. 58 (1), 169–183.
Aura, S., Diamond, P., Geanakoplos, J., 2002. Savings and portfolio choice in a two-period two-asset model. Am. Econ. Rev. 92 (4), 1185–1191.
Bettinger, E., Gurantz, O., Kawano, L., Sacerdote, B., 2019. The long run impacts of merit aid: Evidence from California's cal grant. Am. Econ. J.: Econ. Policy 11 (1), 64–94.
Camerer, C., Babcock, L., Loewenstein, G., Thaler, R., 1997. Labor supply of New York City cabdrivers: One day at a time. Q. J. Econ. 112 (2), 407–441.
Campbell, S., Reeves, S., Kontopantelis, E., Sibbald, B., Roland, M., 2009. Effects of pay for performance on the quality of primary care in England. New Engl. J. Med. 361 (4).
Campos-Mercade, P., 2024. Power Analysis Through Simulations in Stata: A Step-By-Step Guide. Technical Report, Working Paper.
Campos-Mercade, P., Götte, L., Graeber, T.W., Kellogg, A., Sprenger, C., 2024a. "De Gustibus" and Disputes about Reference Dependence. Harvard Business School.
Campos-Mercade, P., Meier, A.N., Meier, S., Pope, D.G., Schneider, F.H., Wengström, E., 2024b. Incentives to Vaccinate. Technical Report, National Bureau of Economic Research.
Campos-Mercade, P., Meier, A.N., Schneider, F.H., Meier, S., Pope, D., Wengström, E., 2021. Monetary incentives increase COVID-19 vaccinations. Science 374 (6569), 879–882.
Campos-Mercade, P., Thiemann, P., Wengstrom, E., 2025. Performance Incentives in Education: The Role of Goal Mismatch. Technical Report, Working Paper.
Campos-Mercade, P., Wengström, E., 2020. Threshold incentives and academic performance. Work. Pap..
Charness, G., Gneezy, U., 2009. Incentives to exercise. Econometrica 77 (3), 909–931.
Cohodes, B.S.R., Goodman, J.S., 2014. Merit aid, college quality, and college completion: Massachusetts' adams scholarship as in in-kind subsidy. Am. Econ. J.: Appl. Econ. 6 (4), 251–285.

Corgnet, B., Gomez-Minambres, J., Hernan-Gonzalez, R., 2015. Goal setting and monetary incentives: When large stakes are not enough. Manag. Sci. 61 (12), 2926–2944.

Corgnet, B., Gómez-Miñambres, J., Hernán-Gonzalez, R., 2018. Goal setting in the principal–agent model: Weak incentives for strong performance. Games Econom. Behav. 109, 311–326.

Cornwell, C.M., Lee, K.H., Mustard, D.B., 2005. Student responses to merit scholarship retention rules. J. Hum. Resour. 40 (4), 895–917.

De Paola, M., Scoppa, V., Nisticò, R., 2012. Monetary incentives and student achievement in a depressed labor market: results from a randomized experiment. J. Hum. Cap. 6 (1), 56–85.

Dechenaux, E., Kovenock, D., Sheremeta, R.M., 2015. A survey of experimental research on contests, all-pay auctions and tournaments. Exp. Econ. 18 (4), 609–669.

Dekel, E., Scotchmer, S., 1999. On the evolution of attitudes towards risk in winner-take-all games. J. Econom. Theory 87 (1), 125–143.

Demski, J., Feltham, G., 1978. Economic incentives in budgetary control systems. Account. Rev..

Eriksen, K.W., Kvaløy, O., 2017. No guts, no glory: An experiment on excessive risk-taking. Rev. Financ. 21 (3), 1327–1351.

Fack, B.G., Grenet, J., 2015. Improving college access and success for low-income students: Evidence from a large need-based grant program. Am. Econ. J.: Appl. Econ. 7 (2), 1–34.

Forsberg, E., Axelsson, B., Arnetz, B., 2001. Financial incentives in health care. The impact of performance-based reimbursement. Health Policy (58), 243–262.

Gill, D., Prowse, V., 2012. A structural analysis of disappointment aversion in a real effort competition. Am. Econ. Rev. 102 (1), 469–503.

Gilpatric, S.M., 2009. Risk taking in contests and the role of carrots and sticks. Econ. Inq. 47 (2), 266–277.

Goerg, S.J., Kube, S., 2012. Goals (th)at work–goals, monetary incentives, and workers' performance. MPI Collect. Goods Prepr. (2012/19).

Gollier, C., 2008. Understanding saving and portfolio choices with predictable changes in assets returns. J. Math. Econom. 44, 445–458.

González-Jiménez, V., 2024. Incentive design for reference-dependent preferences. J. Econ. Behav. Organ. 221, 493–518.

Gravelle, H., Sutton, M., Ma, A., 2010. Doctor behaviour under a pay for performance contract: Treating, cheating and case finding? Econ. J..

Heath, C., Larrick, R.P., Wu, G., 1999. Goals as reference points. Cogn. Psychol. 38 (1), 79–109.

Hvide, H.K., 2002. Tournament rewards and risk taking. J. Labor Econ. 20 (4), 877–898.

Jensen, M., Meckling, W., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. J. Financ. Econ..

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decisions under risk. Econometrica 263–291.

Koch, A.K., Nafziger, J., 2020. Motivational goal bracketing: An experiment. J. Econom. Theory 185, 104949.

Konrad, K.A., et al., 2009. Strategy and dynamics in contests. OUP Cat..

Kräkel, M., Nieken, P., Przemeck, J., 2014. Risk taking and investing in electoral competition. Eur. J. Political Econ. 33, 98–120.

Kremer, M., Miguel, E., Thornton, R., 2009. Incentives to learn. Rev. Econ. Stat. 91 (3), 437–456.

Lacetera, N., Macis, M., Slonim, R., 2013. Economic rewards to motivate blood donations. Science 340 (6135), 927–928.

Lambert, R., Larcker, D., 1991. Executive compensation, corporate decision making, and shareholder wealth: A review of the evidence. In: Executive Compensation. Harvard Business School Press.

Lazear, E.P., Rosen, S., 1981. Rank-order tournaments as optimum labor contracts. J. Political Econ. 89 (5), 841–864.

Leuven, E., Oosterbeek, H., van der Klaauw, B., 2010. The effect of financial rewards on students' achievement: Evidence from a randomised experiment. J. Eur. Econ. Assoc. 8 (6), 1243–1265.

Levitt, S., List, J.A., Sadoff, S., 2017. The effect of performance-based incentives on educational achievement: Evidence from a randomized experiment. In: NBER Working Paper 22107. Work. Pap..

Locke, E.A., Latham, G.P., 1990. A Theory of Goal Setting & Task Performance. Prentice-Hall, Inc.

Milkman, K.L., Gromet, D., Ho, H., Kay, J.S., Lee, T.W., Pandiloski, P., Park, Y., Rai, A., Bazerman, M., Beshears, J., et al., 2021. Megastudies improve the impact of applied behavioural science. Nature 600 (7889), 478–483.

Miller, J., Wiseman, R., Gomez-Mejia, L., 2002. The fit between CEO compensation design and firm risk. Acad. Manag. J. 45 (4).

Montalbán, J., 2018. Countering moral hazard in higher education: The role of performance incentives in.

Murphy, K., 2001. Performance standards in incentive contracts. J. Account. Econ. 30, 245–278.

Nalebuff, B.J., Stiglitz, J.E., 1983. Prizes and incentives: towards a general theory of compensation and competition. Bell J. Econ. 21–43.

Nieken, P., 2010. On the choice of risk and effort in tournaments—experimental evidence. J. Econ. Manag. Strat. 19 (3), 811–840.

Ockenfels, A., Sliwka, D., Werner, P., 2014. Bonus payments and reference point violations. Manag. Sci. 61 (7), 1496–1513.

Palomino, F., Prat, A., 2003. Risk taking and optimal contracts for money managers. RAND J. Econ. 113–137.

Sandmo, A., 1968. Portfolio choice in a theory of saving. Swed. J. Econ. LXX, 106–122.

Sandmo, A., 1970. The effect of uncertainty on saving decisions. Rev. Econ. Stud. 37, 353–360.

Scott-Clayton, J., 2011. On money and motivation a quasi-experimental analysis of financial incentives for college achievement. J. Hum. Resour. 46 (3), 614–646.

Shelev, O., 2012. Incentive thresholds, risk taking, and performance: Evidence from hedge funds. Coleman Fund Risk Manag. Res. Cent. Work. Pap..

Smith, C., Stulz, R., 1985. The determinants of firms' hedging policies. J. Financ. Quant. Econ. 391–405.

Sprinkel, G.B., Williamson, M.G., Upton, D.R., 2008. The effort and risk-taking effects of budget-based contracts. Account. Organ. Soc. 33, 436–452.

Tsetlin, I., Gaba, A., Winkler, R.L., 2004. Strategic choice of variability in multiround contests and contests with handicaps. J. Risk Uncertain. 29 (2), 143–158.

Tullock, G., 1980. Efficient rent seeking. In: Toward a Theory of the Rent-Seeking Society. Texas University Press, College Station, TX, pp. 97–112.

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

Zhou, X., Swan, P., 2003. Performance thresholds in managerial incentive contracts. J. Bus. 76 (4), 665–696.