False Sense of Security? A Study of Risk Compensation in the Lab and the Field*

Lawrence Jin† Gyula Seres‡

October 20, 2023

Abstract

Do safety interventions create a false sense of security? We investigated whether people adjust their behavior optimally when there are exogenous changes to risk. In a laboratory experiment, subjects played an insurance-buying game under various risks of losing endowment. We found strong evidence of overcompensation – subjects purchased more (less) insurance than optimum in response to an increase (decrease) in risk. The degree of overcompensation was larger when risk increased than when it decreased. We exposed the same subjects to changes in safety conditions in real life and found little evidence that lab behaviors predict behaviors in the field.

Keywords: Decision under risk; risk compensation; endowment effect; loss aversion; behavioral economics.

JEL Codes: C91, C93, D9, I12

*We are extremely grateful to Anna Balleyer, Nicola Cerutti, Jana Friedrichsen, Lorenz Goette, Arpita Khanna, Charles Noussair, Rohan Ray, Yohanes Eko Riyanto, Müge Süer, participants of the economics seminar of Nanyang University of Technology and of the 2023 APESA conference, and two anonymous referees for their helpful comments, and Jin Jie Cheow, Jonas Ho, and Maya Kamath for their excellent research assistance. We also thank the Ministry of Education of Singapore for their generous funding that allowed us to conduct this study. This study was approved by the Institutional Review Board of the National University of Singapore (NUS-IRB-2022-158) and preregistered at the AEA RCT registry (AEARCTR-0005831), `https://doi.org/10.1257/rct.9716`. All errors are our own.

†Lee Kuan Yew School of Public Policy, National University of Singapore. Email: lawrencejin@gmail.com.

‡Global Asia Institute, National University of Singapore. Email: sergyu1@gmail.com.

1 Introduction

Do risk-mitigating public policies create a "false sense of security"? Introduced by Peltzman (1975), the risk compensation hypothesis postulates that individual behavioral changes offset the expected positive effects of a public intervention mitigating risk. Peltzman (1975) argued that traffic deaths cannot be successfully prevented by a seat belt mandate as it could induce drivers to compensate by driving more recklessly. That is, individuals would treat the policy and individual effort to mitigate risk as substitutes.

Since the seminal study by Peltzman (1975), the risk compensation hypothesis has been empirically tested in many different contexts. The risk compensation hypothesis has been tested in the context of HIV prevention (Eaton and Kalichman, 2007; Marcus et al., 2013; Wilson et al., 2014), bicycle helmet regulations (Adams and Hillman, 2001), smoking behavior (Evans and Farrelly, 1998; Adda and Cornaglia, 2006; Scherer and Lee, 2014), seat belt legislation (Houston and Richardson, 2007; Evans and Graham, 1991; Cohen and Einav, 2003), face mask mandates (Seres et al., 2021b,a; Mantzari et al., 2020; Howard et al., 2021; WHO, 2020) and vaccines WHO (2021). The empirical results have been mixed, and some studies have questioned whether people compensate risk at all (Pless, 2016; Radun et al., 2018; Mantzari et al., 2020).[1]

Understanding the driving forces behind compensating behavior is crucial in designing risk-mitigating public policies. In particular, the question of whether people compensate risk optimally—as hypothesized by Blomquist (1986)—or whether individuals are prone to making biased adjustments remains unexplored in the literature. In a combined lab and field experiment, we investigated whether people compensate *optimally* in response to exogenous changes in risk. Specifically, we aimed to test whether people have a tendency to *overcompensate*, i.e., reduce (increase) their efforts by more than what is optimal in response to an improvement (deterioration) in safety. The laboratory setup

[1]Several explanations have been proposed to explain why studies have failed to identify compensating behaviors in some settings. Regarding male circumcision as an HIV prevention policy, Wilson et al. (2014) found no evidence of circumcised men engaging in more frequent unprotected sex. They argue that this may be because the policy made more salient the trade-off between engaging in risky behavior and avoiding acquiring HIV. Using an online survey experiment examining the effect of face masks, Seres et al. (2021b) showed that exogenous precautionary behavior by others in a community setting can signal a preference for such behavior.

enabled us to precisely predict optimal risk-compensating behavior, and thus test for risk-overcompensation bias (ROB). We found compelling evidence of ROB – a tendency to adjust more than optimum in response to a change in risk.

In the laboratory experiment, the subjects played five rounds of an insurance-buying game. In each round, they received an endowment that they could lose if an "accident" occurred. They could reduce the probability of an accident by buying costly insurance ("effort") from a set of options. In each round, we exogenously varied the degree of safety, which determined the probability of an accident as well as the cost of insurance. These parameters were carefully designed so that the payoff-maximizing choice remained the same in every round, irrespective of the changes in the safety condition. That is, to maximize payoffs, subjects should buy the same amount of insurance each round.

We found strong evidence of risk-overcompensation bias. On average, subjects bought significantly more (less) insurance when safety deteriorated (improved). We show that risk aversion could not explain the behavior. We estimate that over 70% of the subjects exhibited risk-overcompensation bias.

To test for potential asymmetry in risk-compensating behavior, for one-half of the subjects the external safety condition always improved in each round, and for the other half, the safety condition deteriorated in each round. The rationale was to test for the endowment effect—a well-documented gap between people's willingness-to-pay (WTP) and willingness-to-accept (WTA) in trading goods (Kahneman et al., 1990, 1991).[2] The endowment effect has previously been tested almost exclusively with physical goods. We tested if people also exhibit similar tendencies towards intangibles, like safety. If so, subjects would have a stronger desire to compensate for a loss in safety (i.e., by buying more insurance), than to compensate for a gain in safety (i.e., by buying less insurance).

We observed strong asymmetry in risk overcompensation. The subjects overcompensated for both increases and decreases in risk, but the magnitude of the bias was two times larger when risk increased than when it decreased. The evidence is consistent with

[2]This discrepancy extends to reference dependence effects for increases and decreases in health risk probabilities Viscusi and Huber (2012); Machina and Viscusi (2013). Empirical evidence shows that the difference can be up to two-fold for changes in fatality risk (Viscusi and Evans, 1990).

subjects exhibiting an endowment effect towards safety.

We augmented our laboratory experiment with additional field interventions to test if behaviors in the lab can predict real-life behaviors in the field. That is, are subjects who exhibited a greater degree of risk-overcompensation bias in the lab more likely to exhibit a greater degree of compensating behaviors in the field? Before the laboratory experiment, we exposed the subjects to two health safety field interventions: in half of the sessions, the subjects learned that a high-efficiency particulate absorbing (HEPA) filter was operating in the room; and half of the subjects were required to wear an N95 respirator.[3] We measured their social distancing in a queue, hand sanitizer use, and proper mask use during the sessions. As the subject pool was the same, this design enabled us to compare risk behavior in the lab and in the field. We found that the observed ROB in the laboratory is a weak predictor of risk-compensating behavior in the field.

2 Model and Definitions

Consider an agent facing a risk of experiencing an *accident* – a binary event that occurs with some probability and incurs a loss of utility. The probability of the accident is mitigated by two factors: the safety condition, $w \geq 0$, and the agent's effort to avoid the accident, $e \geq 0$. The safety conditions of the agent are determined exogenously. On the other hand, the agent's effort is determined endogenously; upon observing the safety conditions w, the agent determines an effort level e.

Consider an agent with a von Neumann-Morgenstern utility function that faces a decision-making problem of choosing an effort level $e \geq 0$ under uncertainty to mitigate the risk of an accident. The agent's utility is given by:

$$U(e, w) = p(e, w) \cdot u(I - L - c(e, w)) + (1 - p(e, w)) \cdot u(I - c(e, w)) \tag{1}$$

[3]Note that our experiments took place during the third year of the COVID-19 pandemic. In Singapore at the time, it was mandatory to wear a face mask indoors.

where w is a safety condition observed by the agent; e denotes the agent's effort to mitigate the risk of an accident; I is the agent's endowment; $p(e, w)$ is the probability of an accident, in which case the agent will suffer a loss of L; and $c(e, w)$ is the cost of effort.[4]

We impose assumptions on the probability function $p(e, w)$ and cost function $c(e, w)$. First, both effort e and the safety conditions w decrease the chance of an accident: $\frac{\partial p(e,w)}{\partial e} < 0$ and $\frac{\partial p(e,w)}{\partial w} < 0$. Second, the cost function is increasing and convex in the amount of effort: $\frac{\partial c(e,w)}{\partial e} > 0$ and $\frac{\partial^2 c(e,w)}{\partial^2 e} > 0$.[5]

It is unclear whether the cost function increases, decreases, or remains unchanged with respect to the safety conditions w. This may be context-specific. We believe that in many real-world situations, the cost function is weakly decreasing in w, i.e., $\frac{\partial c(e,w)}{\partial w} \leq 0$. This is because safety interventions are often implemented in tandem with supplementary policies to raise public awareness and make it more convenient for the public to comply.[6]

We denote the agent's choice function as $\bar{e}(w)$ and assume that their unique optimal choice is $e^*(w)$. Given the model and assumptions above, it is ambiguous *ex ante* whether the safety conditions w and effort e are substitutes or complements. On the one hand, when the safety conditions improve, the risk of an accident declines and thus there is an incentive to reduce costly effort. On the other hand, the improved safety conditions could make effort less costly, and thus encourage agents to make more of an effort.

If the safety conditions and effort are substitutes, then the optimal choice of effort $e^*(w)$ is decreasing in w. Formally,

$$\frac{\partial e^*(w)}{\partial w} < 0. \tag{2}$$

We propose two hypotheses, outlined below:

[4]Note that this is a generalization of the framework of Peltzman (1975) whose expected-payoff maximizing model is tailored to road safety. The agent chooses the time devoted to driving and the cost of driving is a linear function of time expressed by forgone income.

[5]Blomquist (1986) explores the objective function $U(e, w) = p(e, w) \cdot u(I - L(e, w) - c(e, w)) + (1 - p(e, w)) \cdot u(I - c(e, w))$. and concludes that the shape of $L(e, w)$ may change the optimal choice of the agent, but this model yields qualitatively similar results to ours.

[6]For example, seat belt beepers installed in cars help drivers remember to comply with seat belt laws, and price subsidies make it cheaper for people to buy masks and comply with mask-wearing regulations.

1. **Risk compensation (RC) hypothesis**: People reduce (increase) effort in response to an improvement (deterioration) in the safety condition. Formally,

$$\frac{\partial \bar{e}(w)}{\partial w} < 0. \tag{3}$$

The existing literature has documented some evidence of risk compensation in the field. However, an unexplored question is whether people compensate for risk optimally. We hypothesize that people may systematically over-respond to changes in the safety conditions:

2. **Risk-overcompensation bias (ROB)**: In response to an improvement (deterioration) in the safety condition, people reduce (increase) effort by more than the optimal amount. Formally,

$$\frac{\partial \bar{e}(w)}{\partial w} < \frac{\partial e^*(w)}{\partial w} \tag{4}$$

The two hypotheses have important implications for the efficacy and welfare consequences of policies aimed at safeguarding the public from harm. Risk compensation, if it occurs, can offset the intended benefits of the safety interventions. Risk-*overcompensation* bias can make situations worse, leading to greater offsets of the intended benefits of the policies. In the field, however, it is difficult to test whether or not an observed level of effort is optimal. Unlike the choice of effort, the cost function is typically not directly observable or quantifiable. Moreover, a decision-maker's beliefs about the chance of an accident also often remain unobserved.

We have designed a laboratory experiment to test for ROB. The underlying idea of the experiment is to present the subjects with a set of choices in the spirit of Equation (1), with each choice featuring a different safety condition.

3 Laboratory Experiment

The laboratory experiment took place at the Center of Behavioral Economics at the National University of Singapore, between July 13 and August 25, 2022.[7] Participants were invited from an online pool of students aged 18 and above. We conducted 32 sessions with 5 to 14 participants, for a total of 314 subjects in the experiment. We used z-Tree (Fischbacher, 2007) to conduct our laboratory sessions. The average session length was 60 minutes. In addition to the earnings from the experiment, the subjects received a show-up fee of S$5.[8]

3.1 Design

Each subject received an initial endowment of S$17.50 and participated in five periods of a decision-making game. Only one randomly chosen period determined their payment. Each period consisted of making choices in two stages. The first stage was a practice simulation for subjects to learn about the safety conditions, but where the subject's choice did not affect payment. Their choice in the second stage counted toward their payoff.

Our laboratory experiment design followed the model in Section 2. In each period, the subject faced a non-zero chance of experiencing an accident in which they would lose S$15 from their endowment. The probability of the accident was in part determined by the safety condition, w, which subjects had no control over. However, the participants could choose a level of insurance, e, to mitigate the risk. Specifically, the probability of an accident was:

$$p(e, w) = \frac{1}{e \cdot w}$$

[7] A detailed experiment protocol is available in the Supplementary Materials. The experiment was pre-registered at the AEA RCT registry: `https://doi.org/10.1257/rct.9716`.

[8] The exchange rate at the time was 0.72 Singapore dollar (S$) = 1 US dollar.

and the cost of the insurance was[9]

$$c(e, w) = \frac{e^2}{w}.$$

The participants made insurance choices across five safety conditions (corresponding to five periods), where w took values $\{4, 5, 6, 8, 10\}$. Subjects could choose any insurance amount $1 \leq e \leq 3.1$ to the nearest one decimal point; an interval that guarantees a non-negative payoff. This intervention tests whether there is a difference between reducing and increasing the safety conditions. Most of the literature on risk compensation entails data on increasing w, but public measures are often temporary, and changes in both directions may trigger ROB.

The subjects learned that in each period, w was an unknown value between 4 and 10.[10] They learned that each period had a different value of w but not that w would be either increasing or decreasing in each period. We allowed subjects to observe and infer the value of w in the practice simulation stage. In this stage, the participants had to enter a permitted e, and the computer displayed the monetary outcome of 10 draws based on the subject's choice of e and the unknown w. The computer screen also showed the number of times the simulation resulted in an accident (i.e., losing S\$15). Subjects then proceeded to the second stage, where they could choose the insurance level e. Only one randomly chosen period determined the subject's final payment.

We used two randomly allotted treatment conditions. In the improving treatment, the five periods increased w from 4 to 10. In the deteriorating treatment, the five periods decreased w from 10 to 4. The purpose of this intervention was to observe whether the adjustment is different between a negative or a positive change in the external safety condition.

[9]Note that the cost function is decreasing in the safety condition, w. As discussed in Section 2, this reflects the fact that many real-world safety interventions are implemented in tandem with supplementary policies to reduce the cost of compliance.

[10]There are at least two reasons for not telling subjects the exact value of w. First, the participants with good computational skills would be able to compute the optimal insurance level e^* if w is known, making the experiment a straightforward mathematical exercise. Second, and more importantly, this approach better resembles many real-life situations where subjects observe the frequency of accidents, but the exact probability of an accident remains unknown.

After five periods of the decision-making game, we collected data on individual risk preferences using a lottery experiment of Holt and Laury (2002). Subjects made 10 binary choices between a “risky” (S\$3.85 or S\$0.10 payment) and a “safe” (S\$2.00 or S\$1.60 payment) gamble. The choices yield an estimate of the coefficient of the constant relative risk aversion (CRRA), henceforth denoted by r.[11] The session ended with a short questionnaire asking for demographic details (see Supplementary Materials).

3.2 Optimal Choice for Risk-Neutral Subjects

The objective function of a risk-neutral subject is:

$$\min_e p(e,w) \cdot L + c(e,w) = \frac{1}{e \cdot w} \cdot 15 + \frac{e^2}{w}.$$

From the first-order condition, we get

$$-\frac{15}{e^2 \cdot w} + \frac{2e}{w} = 0 \iff e^* = 7.5^{\frac{1}{3}} \approx 1.96$$

Hence, the optimal choice $e^*(w)$ is about 1.96, independent of the safety condition w. That is, for subjects trying to maximize the expected payoff from the experiment, it is optimal to not adjust their insurance choice in response to changes in the safety conditions.

An important point we highlight is that the form of the cost function drives the sign of the effect of the external safety condition w on the optimal choice of effort $e^*(w)$. Hence, the result above that the optimal choice is independent of w should not be interpreted as a prediction in general.

To see this, consider that from the implicit function, we have

$$\frac{\partial e^*(w)}{\partial w} < 0 \iff -\frac{\frac{\partial^2 c(e,w)}{\partial e \partial w} + \frac{\partial^2 p(e,w)}{\partial e \partial w} \cdot L}{\frac{\partial^2 c(e,w)}{\partial^2 e} + \frac{\partial^2 p(e,w)}{\partial^2 e} \cdot L} < 0$$

[11] This experiment identifies a range of values for r as the number of choices is discrete. In the calculations, we input the middle of these intervals, except for the extreme cases (only risky or only safe choice) where we use the maximum and minimum values.

if the individual is expected payoff maximizer. If both $p(e,w)$ and $c(e,w)$ are convex functions of e, the sign is determined by the nominator. In reality, the sign of $\frac{\partial^2 c(e,w)}{\partial e \partial w}$ is highly context-dependent. For example, an external safety condition may have no effect on the marginal cost of effort, e.g. seat-belt use does not limit a driver's ability to drive. In other cases, it can decrease the cost of effort, e.g. community mask use may serve as a reminder.

3.3 Evidence of Risk Compensation

We first empirically test whether subjects exhibited risk compensation. When the safety conditions improve, do subjects reduce their insurance choices?

Recall that the subjects did not directly observe the true safety condition, w, but instead observed the frequency of accidents occurring during the practice simulation stage. The probability of an accident equals $p(e,w) = \frac{1}{e \cdot w}$, and subjects observed accident outcomes from 10 random draws. The subjects could infer the most likely underlying safety condition w with one that has the highest probability, $\hat{w}_{i,t} = \frac{10}{e_{i,t}^{sim} \cdot a_{i,t}^{sim}}$, where $e_{i,t}^{sim}$ is the insurance choice made during the practice simulation stage and $a_{i,t}^{sim}$ is the number of accidents observed (between 0 and 10). Since subjects were told that $4 \leq w \leq 10$, $\hat{w}_{i,t} = 4$ if $\frac{10}{e_{i,t}^{sim} \cdot a_{i,t}^{sim}} < 4$ and $\hat{w}_{i,t} = 10$ if $\frac{10}{e_{i,t}^{sim} \cdot a_{i,t}^{sim}} > 10$.

We estimate the following reduced-form random-effects model:

(5) $$e_{i,t} = \beta_0 + \beta_1 \hat{w}_{i,t} + \beta_2 t + \beta_3 \hat{w}_{i,t-1} + \phi X_i + \varepsilon_i.$$

where $e_{i,t}$ is the insurance choice of subject i in period t; $\hat{w}_{i,t}$ is the implied safety condition observed by subject i in period t; $\hat{w}_{i,t-1}$ is its lagged value; and X_i is a vector of control variables that include the subject's estimated coefficient of constant relative risk aversion r_i, gender, and the number of advanced math courses previously taken to control for computational skills.

The coefficient of interest is β_1. $\beta_1 = 0$ would imply that subjects' insurance choices are not influenced by the implied safety condition. This would be consistent with the optimal behavior of payoff-maximizing subjects. $\beta_1 < 0$ ($\beta_1 > 0$) would imply that subjects decrease (increase) the amount of insurance $\bar{e}_{i,t}$ when they observe that safety improves.

The results are shown in Table 1.[12] Column 1 is the simplest model regressing insurance choice ($e_{i,t}$) on the implied safety condition ($\hat{w}_{i,t}$), without any controls. Column 2 controls for time period, t; column 3 further includes the implied safety condition of the previous period; and finally, column 4 includes full set of controls.

Table 1: Main results

	(1)	(2)	(3)	(4)
	$e_{i,t}$	$e_{i,t}$	$e_{i,t}$	$e_{i,t}$
$\hat{w}_{i,t}$	-0.063***	-0.065***	-0.058***	-0.057***
	(0.0078)	(0.0078)	(0.0085)	(0.0086)
t		0.026**	0.023*	0.023*
		(0.0084)	(0.011)	(0.011)
$\hat{w}_{i,t-1}$			-0.028**	-0.027**
			(0.0088)	(0.0089)
r_i				0.026
				(0.065)
Age > 23				-0.050
				(0.071)
Female				-0.0034
				(0.067)
High Math				-0.056
				(0.066)
Constant	2.31***	2.24***	2.36***	2.37***
	(0.048)	(0.052)	(0.077)	(0.098)
Observations	1570	1570	1256	1256
R^2	0.067	0.070	0.078	0.082

Notes: Random-effect GLS regressions. Standard errors in parentheses: † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable $e_{i,t}$ is the amount of insurance purchased by subject i in period t. The main independent variable of interest is $\hat{w}_{i,t}$, the implied safety condition observed by the subject. The CRRA coefficient r_i is estimated with the Holt and Laury (2002) choice experiment. Age > 23 is a dummy variable that equals 1 if the subject's age is above the median age of 23, and zero otherwise. Female is a dummy variable that equals 1 if the subject is female, and zero otherwise. High Math is a dummy variable that equals 1 if the subject took at least 3 university-level math courses, and zero otherwise.

Across all models, the estimated coefficients for the implied safety condition are negative and highly significant. The results suggest that, when safety improved (worsened),

[12]For an overview of the demographic variables, see Table A1.

subjects compensated by buying less (more) insurance. The results provide compelling evidence of risk compensation. The estimated coefficients range from -0.057 to -0.065, implying that when the implied safety condition improved by one unit, on average, subjects reduced their insurance purchase by about 0.06 units. Note that this response was costly and reduced the subject's expected payoff, as shown in Section 3.2.

Interestingly, the estimated coefficient for time period t is positive and significant at the 5% significance level. This suggests that subjects, on average, bought more insurance over time. The estimated coefficients for lagged implied safety condition, $\hat{w}_{i,t-1}$, are negative and significant at the 1% significance level. The magnitudes of the coefficients are about half of the coefficients for the current implied safety conditions. The results suggest that the implied safety condition of the previous time period continued to influence the subjects' insurance purchase decisions, albeit at smaller magnitudes. To account for heterogeneity in analytical skills, we control for self-reported mathematical skills but find no evidence for an effect.

Note that in our experiment, the subjects did not directly observe the safety condition, but instead could infer the safety condition from the practice simulation results (i.e., display of outcomes from ten random draws). This was implemented to better resemble many real-world situations where subjects observe the frequency of accidents, but the exact probability of an accident remains unknown. As a robustness check, we estimated the same models with the true underlying value of safety condition ($w_{i,t}$) as the main independent variable, and found consistent results (Table A2 in the Supplementary Materials).

In Table 2, we conducted heterogeneity analyses to examine whether some subjects tended to exhibit a greater or weaker degree of risk compensation. In column (1), we included an interaction term between the implied safety condition and an indicator for whether the subject is female. The estimated coefficient for the interaction term is negative and significant at the 5% level. On average, female subjects exhibited 83% greater degree of risk compensation than male subjects (-0.042 for males, -0.077 for females). In column (2), we included an interaction term between the implied safety condition and whether the subject has taken three or more university-level mathematics courses. The estimated coefficient of the interaction is positive and statistically significant at the 5% level. The results suggest that subjects who have taken more mathematics courses showed a lesser degree of risk compensation. Finally, in column (3), we tested for an interaction effect between implied safety condition and an indicator for the subject being older than the median age of 23 years of age. We found no significant difference in the degree of risk compensation exhibited by younger vs. older subjects.

Table 2: Heterogeneity analyses

	(1)	(2)	(3)
	$e_{i,t}$	$e_{i,t}$	$e_{i,t}$
$\hat{w}_{i,t}$	-0.042***	-0.074***	-0.063***
	(0.012)	(0.011)	(0.0099)
$\hat{w}_{i,t}$ × Female	-0.035*		
	(0.017)		
$\hat{w}_{i,t}$ × High Math		0.043*	
		(0.017)	
$\hat{w}_{i,t}$× Age > 23			0.015
			(0.013)
$\hat{w}_{i,t-1}$	-0.028**	-0.027**	-0.027**
	(0.0089)	(0.0088)	(0.0088)
t	0.025*	0.023*	0.023*
	(0.011)	(0.011)	(0.011)
r_i	0.026	0.024	0.016
	(0.064)	(0.065)	(0.065)
Age > 23	-0.012	-0.012	-0.018+
	(0.0087)	(0.0089)	(0.010)
Female	0.17	-0.0039	0.015
	(0.10)	(0.063)	(0.065)
High Math	-0.039	-0.26*	-0.035
	(0.066)	(0.11)	(0.067)
Constant	2.55***	2.72***	2.78***
	(0.23)	(0.23)	(0.26)
Observations	1256	1256	1256
R^2	0.090	0.085	0.079

Notes: Random-effect GLS regressions. Standard errors in parentheses: † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable $e_{i,t}$ is the amount of insurance purchased by subject i in period t. The main independent variable of interest is $\hat{w}_{i,t}$, the implied safety condition observed by the subject. The CRRA coefficient r_i is estimated with the Holt and Laury (2002) choice experiment. Age > 23 is a dummy variable that equals 1 if the subject's age is above the median age of 23, and zero otherwise. Female is a dummy variable that equals 1 if the subject is female, and zero otherwise. High Math is a dummy variable that equals 1 if the subject took at least 3 university-level math courses, and zero otherwise.

3.4 Optimal Choice for CRRA Subjects

Can risk aversion explain why subjects compensated for changes in risk at the cost of their expected payoffs? Let us assume that subjects' preferences exhibit constant relative risk aversion (CRRA) with coefficient r, and their isoelastic utility function is given by $u(c) = \frac{c^{1-r}-1}{1-r}$. Their objective function becomes:

$$\text{(6)} \qquad \max_{e} U(e,w,r) = (1 - \frac{1}{e \cdot w}) \cdot \frac{(17.5 - \frac{e^2}{w})^{1-r} - 1}{1-r} + \frac{1}{e \cdot w} \cdot \frac{(2.5 - \frac{e^2}{w})^{1-r} - 1}{1-r}.$$

where 17.5 and 2.5 are the payoffs with and without an accident, respectively.

Solving this optimization problem computationally, we find that for risk-averse subjects the optimal insurance choice in fact *increases* with improved safety conditions (i.e., if $r > 0$ then $\frac{\partial e^*(w)}{\partial w} < 0$). Risk aversion cannot explain why subjects compensated risk in our laboratory experiment. On the other hand, for risk-loving subjects the optimal insurance choice decreases with improved safety conditions. For risk-neutral subjects, the optimal choice is independent of w, as shown in Section 3.2. Given that most of our subjects were risk averse (e.g., see Table A1), risk compensation was unlikely to be utility-maximizing.

3.5 Evidence of Risk-Overcompensation Bias

Next, we test for the risk-overcompensation bias (ROB). For subject i, we denote the optimal choice vector as $e^*_{i,t}(w_{i,t}, r_i)$ calculated from the objective function (6), where $w_{i,t} \in \{4, 5, 6, 8, 10\}$.[13] Similarly, the observed choice vector is $\bar{e}_{i,t}(w_{i,t}, r_i)$. For both, we estimate the slopes β^*_i and $\bar{\beta}_i$ of the linear models:

$$e^*_{i,t} = \alpha^*_i + \beta^*_i w_{i,t} + \varepsilon_{i,t}$$

and

$$\bar{e}_{i,t} = \bar{\alpha}_i + \bar{\beta}_{i,t} w + \varepsilon_{i,t}$$

[13]Point predictions of the optimal choices were calculated with Wolfram Mathematica.

individually for all subjects. We define β_i^* and $\bar{\beta}_i$ as optimal and observed degree of risk compensation for subject i, respectively. Note that these estimates capture the average adjustment along the range of safety conditions.

We found strong support for ROB. The mean optimal degree of risk compensation β_i^* in the sample was 0.027. That is, when we account for the subjects' degree of risk aversion, it would have been optimal for subjects to *increase* their insurance in response to an improvement in the safety conditions. In sharp contrast, the mean observed degree of risk compensation $\bar{\beta}_i$ was negative, at -0.034. The difference between optimal and observed estimates were highly significant (two-sample t-test, t=-20.178, $p < 0.001$). The observed adjustment behavior exhibited standard deviations that were more than four times larger (SD=0.117) than the optimum adjustment behavior (SD=0.027). 71.6% of the subjects' observed slope were smaller than the optimal slope.

Next, we examined asymmetry in risk compensating tendencies between scenarios where the safety condition is improving vs. deteriorating. The mean of observed degree of risk compensation, $\bar{\beta}_i$, was $-.021$ (one-sample t-test vs. 0, $t = -2.3621$, $p = 0.0097$) in the improving scenario and -0.050 (one-sample t-test vs. 0, $t = -5.1081$, $p = 0.0000$) in the deteriorating scenario, respectively. The difference between the two slopes was statistically significant (two-sample t-test, $t = 2.196$, $p = 0.014$). Fig 1 plots the cumulative distribution functions of the observed degree of risk compensation ($\bar{\beta}_i$) between improving (blue) and deteriorating (red) safety conditions. In addition to a significant difference in the mean, there is first-order stochastic dominance between the two conditions.

The results suggest two facts. First, the subjects exhibited significantly stronger degree of risk-compensating tendencies when the safety condition is deteriorating than when it was improving. This is consistent with subjects exhibiting an endowment effect towards safety, i.e., subjects exhibiting stronger desire to compensate when they are about to lose a degree of safety than when they are about to gain it. Second, the degree of risk compensation in the improving-safety scenario was smaller in magnitude but still significantly different from zero. This suggests that the observed risk-compensating tendencies cannot be entirely explained by the endowment effect. That is, there appears

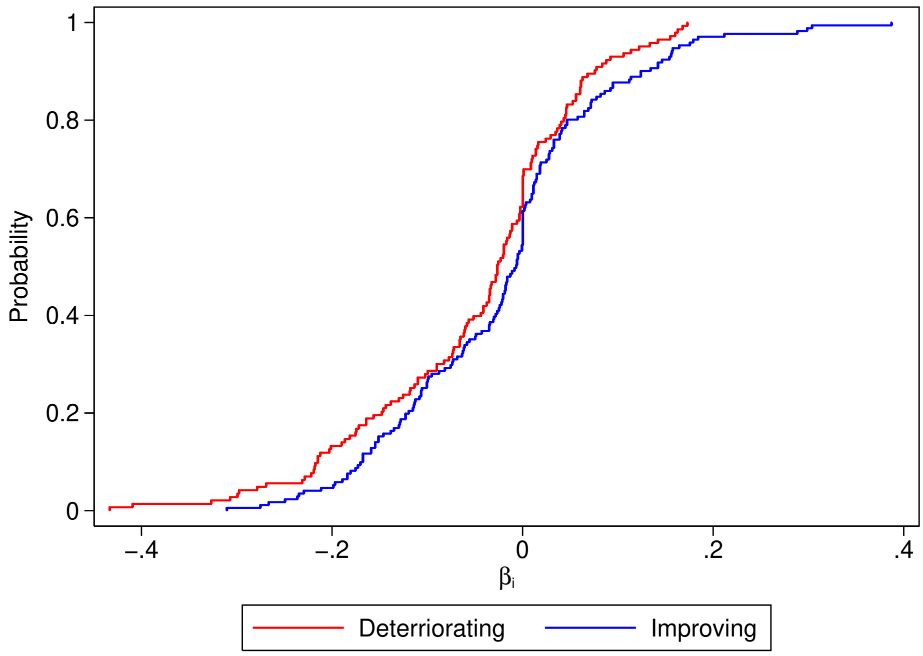

Figure 1: Cumulative distribution functions of the observed degree of risk compensation ($\bar{\beta}_i$) in improving (blue) and deteriorating (red) safety conditions.

to be an underlying risk-overcompensation bias, which is amplified by the endowment effect when safety is deteriorating.

4 The Field Experiment

In our study, we also examined whether the subjects would engage in risk-compensating behaviors in more realistic situations involving non-monetary risk. Specifically, our goal was two-fold: (a) to examine whether we would observe a similar degree of risk-compensating behavior in a field setting; and more importantly, (b) whether there is strong individual-level correspondence between one's risk-compensating tendencies in the lab and in the field.

4.1 Design

We introduced two field safety interventions prior to conducting the laboratory experiment. Both the field and the laboratory experiments took place during the same sessions with the same n=314 subjects. We utilized two rooms: a *waiting room* and the *labora-*

tory. Prior to the laboratory experiment, subjects entered the waiting room for on-site registration. We tested whether the subjects exhibited risk-compensating behavior in response to two safety interventions below.[14] Both interventions were designed to improve the subjects' perceived risk of catching the disease.[15]

1. **N95 Mask Intervention:** Randomized at the individual level, 50% of subjects were notified prior to the session that they will have to wear an N95 or equivalent respirator.[16]

2. **HEPA Intervention:** In half of the sessions, the subjects learned that an air purifier equipped with a HEPA filter was operating in waiting the room. This information together with a government-supported claim that it may destroy up to 99.7% of airborne germs was displayed on a wide-screen LED TV and on flyers. In the control sessions, although the same air purifier was operating, there was no such information on display.

These interventions change the external safety conditions of the subjects, but have different interpretations. With the HEPA Intervention, the baseline is a setting in which there is no intervention. The N95 Mask Intervention's effect shall be interpreted differently. There, the baseline is wearing any mask due to the prevailing mask mandate of the building while knowing that others wear a higher-grade respirator. This is nevertheless an external safety condition as the subjects know that half of the room is subjected to a stricter mandate, and wearing an N95 is not their choice. Hence, wearing it does not signal a preference.[17]

The first outcome variable was the distance each subject stood from the person in front of them in a queue. At the on-site registration in the waiting room, each subject received a randomly drawn identification number. An experimenter in the room then called them

[14] For technical details, please find the Protocol in the Supplementary Materials.

[15] The N95 mask intervention improves both perceived as well as actual risk; the HEPA intervention only improves the perceived risk, since the same HEPA filter was operating in all conditions.

[16] Those who did not bring a required respirator received one free of charge. At the time of the sessions, it was mandatory in Singapore to wear a face mask in all indoor areas, and most people wore either disposable surgical masks or fabric masks.

[17] Seres et al. (2021b) provides incentivized survey evidence for a signaling effect of masks if wearing one is voluntary.

up based on the order of the identification number and asked them to form a queue to enter the laboratory. While the subjects were being sent to the laboratory one by one, an experimenter measured the distance between the first and second subjects in the queue to the nearest 1 centimeter. This distance was assigned as an outcome variable for the second subject in the queue.[18] We recorded two more outcome variables associated with precautionary behavior during the laboratory experiment. In the laboratory, we provided each subject with hand sanitizer wipes at their workstation and recorded whether the subjects used it. We also recorded the number of times each subject removed their face mask during the experiment session, e.g., to drink water.

4.2 Empirical Specification

Both the mask and the HEPA interventions are expected to improve the subjects' perceived safety conditions related to COVID-19.[19] We wanted to examine whether, as a result, participants compensate for the change in perceived risk by being less likely to engage in precautionary behaviors.

For two outcome variables—the queuing distance (DISTANCE) and the number of times the subject pulled down the mask during the session (MASKOFF)—we estimate the following OLS:

$$y_i = \gamma_0 + \gamma_1 MASK_i + \gamma_2 HEPA_i + \gamma_3 MASK_i * HEPA_i + \phi X_i + \varepsilon_i. \tag{7}$$

where y_i is the outcome variable; $MASK_i$ and $HEPA_i$ are indicators for receiving the mask and HEPA interventions, respectively; and X_i is a vector of subject-specific control

[18] This measurement is possible for everyone in the room except the first subject in line. To reduce the loss of data, most sessions had a confederate who took the very first position in the queue. Some sessions did not have a confederate. In this case, no distance was recorded for the first subject in the waiting line.

[19] Subjects were notified prior to the experiment whether they had been selected into the masking treatment. This could potentially create a selection bias. We tested the independence with the following logit binary choice model: $Pr(MaskS = 1) = \frac{\exp(\gamma+\phi X_i+\varepsilon_i)}{1+\exp(\gamma+\phi X_i+\varepsilon_i)}$ (n=314) in which X_i is the vector of personal characteristics (age, gender, field of study, number of mathematics courses). No coefficient is significant at the 5% level.

variables. For the third outcome variable—the use of hand sanitizer wipes (Sanitizer)—we estimate an equivalent logit model. The standard errors are clustered at the session level as the behaviors of the other subjects in the session are (partially) observable.

4.3 Results

The results are shown in columns (1)–(3) of Table 3.[20] Overall, we find mixed evidence of risk-compensating behavior in the field. We find that wearing a higher-grade protective mask led to a decrease in the use of hand sanitizer wipes (coefficient = -0.74, $p = 0.021$), but there were no significant effects on either the queuing distance or the frequency of taking off masks. A HEPA intervention may have led to an increase in the frequency of mask removal, but the estimated effect was only marginally significant at the 10% level (coefficient = 0.51, $p = 0.070$). HEPA intervention had no significant effect on either the queuing distance or the use of hand sanitizer. None of the interaction terms were statistically significant.[21]

The fact that we do not find robust evidence of risk compensating behavior in the field is broadly in line with the empirical literature, which has been largely mixed. A plausible reason behind this is that the cost function of effort $c(e, w)$ may have different properties in different settings. Depending on how a change in the safety condition affects the marginal cost of effort, it may or may not be optimal for subjects to compensate for risk at all.[22] Specifically, if an improvement in the safety condition sufficiently reduces the marginal cost of effort, w and e may in fact be complements.

Next, we test whether subjects who have exhibited greater risk-compensating tendencies in the lab were also more likely to compensate for risk in the field. In columns (4)–(6) of Table 3, we augment the regression models by including the observed individual degree of risk compensation in the lab, $\bar{\beta}_i$, as well as its interactions with the two field interven-

[20] The main descriptive statistics are summarized in Table A1 in the Supplementary Materials.

[21] Note that the regressions only include three of the nine questions eliciting the subject's opinions about preventive measures. The other six measures (regular testing, drinking hot water) are not related to the outcome variables.

[22] The most controversial assumption of Blomquist (1986) is that the safety conditions do not decrease the marginal cost of effort: $\frac{\partial \partial c(e,w)}{\partial e \partial w} \geq 0$. If this does not hold, we may have $\frac{\partial e^*(w)}{\partial w} < 0$ and no risk compensation (Seres et al., 2021a).

Table 3: The Effect of Interventions on Precautionary Behavior

	(1) Distance	(2) Sanitizer	(3) MaskOff	(4) Distance	(5) Sanitizer	(6) MaskOff
Treatment HEPA	4.08	0.66	0.51^{+}	4.64	0.92	0.63*
	(4.84)	(0.64)	(0.27)	(5.03)	(0.62)	(0.29)
Treatment MASK	1.10	-0.74*	0.043	0.75	-0.56^{+}	-0.0042
	(3.46)	(0.32)	(0.15)	(3.32)	(0.31)	(0.15)
HEPA $\times$ MASK	-7.97	0.29	-0.44	-8.84^{+}	0.046	-0.53
	(4.74)	(0.50)	(0.33)	(4.81)	(0.52)	(0.33)
$\bar{\beta}_i$				26.6	5.14	1.05
				(21.9)	(4.20)	(0.97)
HEPA $\times$ $\bar{\beta}_i$				-13.2	-2.12	1.25
				(24.7)	(4.22)	(1.07)
MASK $\times$ $\bar{\beta}_i$				-28.0	-2.96	-2.50^{+}
				(23.3)	(3.25)	(1.35)
Queue Length	-6.60***			-6.73***		
	(1.22)			(1.23)		
Female	0.75	0.025	-0.50*	0.80	0.029	-0.52**
	(2.69)	(0.48)	(0.18)	(2.73)	(0.48)	(0.19)
Importance Indoor Masking	0.59	-0.086	-0.11	0.65	-0.13	-0.11
	(1.39)	(0.14)	(0.10)	(1.39)	(0.18)	(0.11)
Importance Ventillation	1.42	0.16	0.084	1.44	0.26	0.10
	(1.34)	(0.14)	(0.065)	(1.41)	(0.17)	(0.063)
Importance Air Filtering	-0.86	-0.043	-0.019	-1.05	-0.12	-0.030
	(1.40)	(0.14)	(0.074)	(1.45)	(0.12)	(0.071)
Age	0.41	-0.11^{+}	0.020	0.37	-0.11*	0.015
	(0.55)	(0.058)	(0.023)	(0.53)	(0.051)	(0.021)
Study Master	5.78	0.71	-0.51*	5.99	0.45	-0.53*
	(6.15)	(0.79)	(0.19)	(6.23)	(0.91)	(0.19)
Study PhD	11.0^{+}	0.13	-0.30	11.1^{+}	0.054	-0.28
	(5.78)	(1.24)	(0.33)	(5.49)	(1.19)	(0.34)
Comfort Chair		0.22	0.080			
		(0.24)	(0.074)			
Comfort Mask		-0.30*	-0.036			
		(0.14)	(0.057)			
Comfort Temperature		0.46*	-0.0015			
		(0.21)	(0.076)			
Constant	63.2***	-2.55	0.56	65.3***	0.42	1.06
	(12.4)	(1.71)	(0.99)	(12.3)	(1.55)	(0.72)
Observations	304	314	314	304	314	314
R^2	0.1597		0.0845	0.1650		0.0964
Pseudo R^2		0.0987			0.0712	

Notes: OLS (columns 1, 3, 4, 6) and logistic regression (columns 2, 5) estimates. Standard errors are clustered at the session level and are in parentheses. † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. In half of the sessions, subjects were told that a high-efficiency particulate absorbing (HEPA) filter was operating (HEPA). One half of the subjects were required to wear an N95 or equivalent respirator throughout the session (MASK). DISTANCE is the distance between the subject and the person in front of them in a queue, before entering the laboratory. MASKOFF is the number of times the subject removed their mask during the session. SANITIZER=1 if the subject used or took the sanitizer wipe, and 0 otherwise. The questionnaire elicited the subjects' opinions on a 7-point Likert scale.

tions. The estimated coefficients for the interactions reveal whether the individuals who exhibited a greater degree of risk compensation in the lab were also more responsive to the field interventions.

Overall, we do not find such evidence. None of the estimated coefficients for the interactions were statistically significant at the 5% level. The estimated coefficient for MASK $\times\bar{\beta}_i$ was negative and marginally significant at the 10% significance level. This estimate implies that, on average, subjects who exhibited a greater degree of risk compensation in the lab (i.e., lower $\bar{\beta}_i$) were more likely to take off the mask in response to the mask intervention.

5 Discussion

The concern that risk-mitigating measures may create a false sense of security—the risk compensation hypothesis—is a recurring point in public debates. The bleak implication of this hypothesis is that reduced individual effort will offset the intended benefits of policy interventions in crisis situations. Previous research on risk compensation using field and observational data has been largely mixed and inconclusive. We address this puzzle by formally defining risk compensation (RC), and testing it together with the hypothesis of risk-overcompensation bias (ROB).

One major challenge of studying risk compensation in the field has been that people's cost of effort are not observable, which makes it impossible to determine what their optimal response to a safety intervention is. We overcome this problem in our laboratory experiment by precisely defining the cost of effort (i.e., insurance). In the experiment, subjects faced risks of losing their initial endowment, which could be mitigated by buying costly insurance. We exposed the subjects to varying degree of risks and observed their insurance choices. Subjects exhibited a strong tendency to overreact to changes in risk, exhibiting ROB. We found that over 70% of the subjects exhibited ROB. We also found strong asymmetry in their behavior. The subjects exhibited significantly larger degree of risk compensation when safety was deteriorating than when it was improving. The evidence is consistent with subjects exhibiting an endowment effect towards safety.

To test whether people's ROB in the lab predicts their tendency to compensate for risk in the field, we exposed the subjects of the laboratory experiment to two real-world safety interventions. Randomized at the session level, half of the subjects learned about a HEPA filter operating in the room; and randomized at the individual level, half the subjects were instructed to wear a more protective face mask. We examined whether the subjects exposed to the interventions were less likely to engage in precautionary measures: standing closer in the queue, being less likely to use sanitizer wipes, and removing their mask more frequently. We found suggestive but weak evidence of risk compensation in the field. Interestingly, we found no link between the lab and the field: The degree of risk-compensating behavior in the laboratory experiment did not predict the degree of precautionary behavior in the field.

The lack of explanatory power of behaviors in the lab in predicting behaviors in the field calls for caution, on at least two fronts. First, our results illustrate the difficulty of precisely identifying risk compensation in the field. In the field, there is significant heterogeneity across individuals in both the perceived changes in risk and the cost of effort in response to safety interventions. This can explain why previous empirical evidence on risk compensation has been largely mixed and highlights the need to study risk compensation in more controlled environments. Second, risk-compensating tendencies may be highly context-specific.

Regarding the underlying cognitive process, we do not believe that ROB is an error in probability reasoning. In the lab, the subjects received an unbiased estimate of the probability of an accident.[23] One possible cause of ROB may stem from inattention. For example, people have a tendency to pay disproportionately more attention to information that is salient to them, and little attention to other information (e.g., Chetty et al. (2009); Finkelstein (2009); Bordalo et al. (2013)). In our experiment set-up, when safety

[23]For example, consider the generally accepted theory of logarithmic perceived probabilities (Gabaix, 2019). As Steiner and Stewart (2016) point out, individuals may underestimate high and overestimate low probabilities. In our experiment, the chance of an accident is always low, never higher than 25%. An increase in the safety condition w reduces the probability of an accident. An overestimation of the (low) probability of an accident would make subjects want to buy *more* insurance than optimal, not less. Hence, logarithmic perceived probabilities cannot explain ROB.

improves, there are two forces in effect: on the one hand, the individual would want to reduce effort because there is lower risk, but on the other hand, the individual would want to increase effort because the effort is less costly. If people pay too much attention to reduced risk and too little attention to reduced cost of effort, they could end up reducing effort by more than what would be optimal. Future studies could explore the underlying mechanisms behind this result.

We also found that subjects overcompensated more when safety was deteriorating than when it was improving. This is likely linked to the endowment effect (Kahneman et al., 1990, 1991), specifically, the fact that people's willingness-to-accept values for increases in risk tend to be higher than the willingness-to-pay values for risk decreases (Machina and Viscusi, 2013). If this gap holds in general, implementing safety-improving measures in a crisis situation and then later removing them (or vice versa) can lead to long-term effects on community behavior. Hence, understanding the underlying mechanism and its possible link to the endowment effect would be crucial for long-term public risk management.

References

Adams, John and Mayer Hillman, "The risk compensation theory and bicycle helmets," *Injury Prevention*, 2001, *7* (2), 89–91.

Adda, Jerome and Francesca Cornaglia, "Taxes, cigarette consumption, and smoking intensity," *American Economic Review*, 2006, *96* (4), 1013–1028.

Blomquist, Glenn, "A utility maximization model of driver traffic safety behavior," *Accident Analysis & Prevention*, 1986, *18* (5), 371–375.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and consumer choice," *Journal of Political Economy*, 2013, *121* (5), 803–843.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and taxation: Theory and evidence," *American Economic Review*, 2009, *99* (4), 1145–77.

Cohen, Alma and Liran Einav, "The effects of mandatory seat belt laws on driving behavior and traffic fatalities," *Review of Economics and Statistics*, 2003, *85* (4), 828–843.

Eaton, Lisa A and Seth C Kalichman, "Risk compensation in HIV prevention: implications for vaccines, microbicides, and other biomedical HIV prevention technologies," *Current HIV/Aids Reports*, 2007, *4* (4), 165–172.

Evans, William N. and John D. Graham, "Risk reduction or risk compensation? The case of mandatory safety-belt use laws," *Journal of Risk and Uncertainty*, 1991, *4* (1), 61–73.

Evans, William N and Matthew C Farrelly, "The compensating behavior of smokers: taxes, tar, and nicotine," *The Rand Journal of Economics*, 1998, pp. 578–595.

Finkelstein, Amy, "E-ztax: Tax salience and tax rates," *The Quarterly Journal of Economics*, 2009, *124* (3), 969–1010.

Fischbacher, Urs, "z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, 2007, *10* (2), 171–178.

Gabaix, Xavier, "Behavioral inattention," in "Handbook of Behavioral Economics: Applications and Foundations 1," Vol. 2, Elsevier, 2019, pp. 261–343.

Holt, Charles A and Susan K Laury, "Risk aversion and incentive effects," *American Economic Review*, 2002, *92* (5), 1644–1655.

Houston, David J. and Lilliard E. Richardson, "Risk compensation or risk reduction? Seatbelts, state laws, and traffic fatalities," *Social Science Quarterly*, 2007, *88* (4), 913–936.

Howard, Jeremy, Austin Huang, Zhiyuan Li, Zeynep Tufekci, Vladimir Zdimal, Helene-Mari van der Westhuizen, Arne von Delft, Amy Price, Lex Fridman, Lei-Han Tang et al., "An evidence review of face masks against COVID-19," *Proceedings of the National Academy of Sciences*, 2021, *118* (4).

Kahneman, Daniel, Jack L Knetsch, and Richard H Thaler, "Experimental tests of the endowment effect and the Coase theorem," *Journal of political Economy*, 1990, *98* (6), 1325–1348.

_ , _ , and _ , "Anomalies: The endowment effect, loss aversion, and status quo bias," *Journal of Economic perspectives*, 1991, *5* (1), 193–206.

Machina, Mark and Kip Viscusi, *Handbook of the Economics of Risk and Uncertainty*, Newnes, 2013.

Mantzari, Eleni, G James Rubin, and Theresa M Marteau, "Is risk compensation threatening public health in the COVID-19 pandemic?," *BMJ*, 2020, *370.*

Marcus, Julia L., David V. Glidden, Kenneth H. Mayer, Albert Y. Liu, Susan P. Buchbinder, K. Rivet Amico, Vanessa McMahan, Esper Georges Kallas, Orlando Montoya-Herrera, Jose Pilotto et al., "No evidence of sexual risk compensation in the iPrEx trial of daily oral HIV preexposure prophylaxis," *PloS One*, 2013, *8* (12).

Peltzman, Sam, "The effects of automobile safety regulation," *Journal of Political Economy*, 1975, *83* (4), 677–725.

Pless, Barry, "Risk compensation: Revisited and rebutted," *Safety*, 2016, *2* (3), 16.

Radun, Igor, Jenni Radun, Mahsa Esmaeilikia, and Timo Lajunen, "Risk compensation and bicycle helmets: A false conclusion and uncritical citations," *Transportation Research Part F: Traffic Psychology and Behaviour*, 2018, *58*, 548–555.

Scherer, Gerhard and Peter N Lee, "Smoking behaviour and compensation: a review of the literature with meta-analysis," *Regulatory Toxicology and Pharmacology*, 2014, *70* (3), 615–628.

Seres, Gyula, Anna Balleyer, Nicola Cerutti, Jana Friedrichsen, and Müge Süer, "Face mask use and physical distancing before and after mandatory masking: No

Evidence on risk compensation in public waiting lines," *Journal of Economic Behavior & Organization*, 2021, *192*, 765–781.

_ , **Anna Helen Balleyer, Nicola Cerutti, Anastasia Danilov, Jana Friedrichsen, Yiming Liu, and Müge Süer**, "Face masks increase compliance with physical distancing recommendations during the COVID-19 pandemic," *Journal of the Economic Science Association*, 2021, pp. 1–20.

Steiner, Jakub and Colin Stewart, "Perceiving prospects properly," *American Economic Review*, 2016, *106* (7), 1601–31.

Viscusi, W Kip and Joel Huber, "Reference-dependent valuations of risk: Why willingness-to-accept exceeds willingness-to-pay," *Journal of Risk and Uncertainty*, 2012, *44* (1), 19–44.

_ **and William N Evans**, "Utility functions that depend on health status: estimates and economic implications," *The American Economic Review*, 1990, pp. 353–374.

WHO, "Advice on the use of masks in the context of COVID-19," 2020. WHO reference number: WHO/2019-nCoV/IPC_Masks/2020.3.

_ , "WHO Director-General's opening remarks at the media briefing on COVID-19 - 24 November 2021 ," 2021. https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19—24-november-2021.

Wilson, Nicholas L., Wentao Xiong, and Christine L. Mattson, "Is sex like driving? HIV prevention and risk compensation," *Journal of Development Economics*, 2014, *106*, 78–91.

Worldometers, "COVID-19 Coronavirus Pandemic Case Numbers," 2022.

Supplementary Materials A: Experiment Protocol

This section describes the protocol for the decision-making experiment. It was given to the four experimenters that carried out the sessions. The sessions took place from July 13 to August 25, 2022, at the Centre for Behavioural Economics, at the National University of Singapore, at a time when the number of COVID-19 cases in Singapore had reached a record high and ranged between 1.36–1.85% of the resident population (Worldometers, 2022).

Outline

This is a combined field and laboratory experiment that takes place in two neighboring rooms of the Innovation 4.0 building of the National University of Singapore (NUS). One is Meadow 2 (henceforth *waiting room*), the other is the Center for Behavioral Economics (henceforth *laboratory*). The participants are students of the NUS that registered for the study as subjects. Before the session, you will receive the list of participants on which, those that are marked received a notification letter 24 hours prior to their session that they are required to wear a higher grade face mask.[24]

The experiment has two interventions. 1) In half of the sessions, a tv screen mounted on the wall informs the participants that there is an air purifier working in the room with a HEPA filter. The purifier is always switched on, but there is no display in the baseline treatment. The treatment condition for the first session is determined by a coin toss, then the two alternate between sessions (there will be an even number of sessions). 2) Half of the participants will be asked to wear a KF94 (or equivalent or higher grade) respirator during the entire experiment. This will be randomized as follows. The registration deadline is 24 hours before the scheduled experiment. Then, putting the participants in alphabetic order and determined by a coin toss, the first person will be either in the treatment group or not, odd numbers will be assigned to this condition and even number to the other one. Those selected will receive a notification after the registration deadline

[24]The notification explains that these mask types are accepted as higher grade: N95, KF94, FFP2, or higher. The general inclusion criteria is that the face covering must be certified to filter at least 90% of microscopic particles that may carry viruses or bacteria.

has passed (Appendix 2 of this document).

The subjects can enter the waiting room when they arrive. In this room, there are three experimenters, numbered according to their role as E1-E3. Upon entering the room, the subjects start at the registration desk. There, Experimenter 1 records their arrival, gives them a randomized identification number (ID) that is between 2-13, and a sticker with this number. E2 plays the role of a participant and always receives ID 1. E1 i) records the subjects' arrival, ii) gives a sticker to them with the ID number and asks them to put it on their shoulder, iii) gives instructions regarding the general rules of the experiment (e.g. no communication and no phones are allowed), iv) hands over a questionnaire (Appendix 1) and tells them to give it in later in the laboratory, and v) takes care of the masking treatment. For subjects selected for wearing an KF94 respirator, the following protocol applies. First, E1, who has a complete list of participants and their treatment group, asks them whether they have received the notification. Then, E1 confirms their status and asks them what sort of mask they are wearing.[25] After clarifying this, E1 states this: "Higher grade masks with better filtration efficiency protect you better from COVID-19".[26].If they are not wearing a required mask type, E1 offers them a KF94 respirator free of charge and asks them to put it on. E1 fills out these parts of the questionnaire: Date and time and ID number; and supervises the answers to the first questions about the notification email, receiving a mask in the waiting room, and the mask they are wearing.

After registration, the subjects are asked to take a seat and fill out the rest of the questionnaire. Then, they are called to form a waiting line in front of the door in the order of their ID number. There are always maximum six people in the waiting line. E1 coordinates this by first asking ID 1-6 to line up, then asks one person to leave the room and the next person to join the line at each turn. The place for the first person is marked by a tape on the floor.

E3 is responsible for measuring the distancing between subjects as follows. E3 uses a smartphone with an installed augmented-reality tape-measure app that is capable of

[25] E1 is familiar with all publicly available mask types. All subjects are required to wear some sort of mask in all indoor settings according to the prevailing campus regulations.

[26] This corresponds to the guidelines of the Ministry of Health https://www.moh.gov.sg/covid-19/general/faqs—masks-and-personal-protective-equipment-(ppe)

measuring small distances in centimeters (Measure app, iPhone 12). The app measures distance by pinning two points on the ground. These two points are pinned to the closest points of the first and second subject's shoes. The measurement should commence after these two people stop moving. The distance kept between them is the outcome variable for the second person. The queuing order makes E2 the first in the line, enabling the measurement of all subjects. E3 must make sure that the measurement is not obvious to the participants. By doing so, E2 sits perpendicular to the waiting line facing the space between the first two people. E1 monitors this and calls the subjects one-by-one to leave the room and go to the laboratory. The waiting room stage ends when every subjects left the room.

In the laboratory, another experimenter (E4) welcomes the subjects who arrive one-by-one. E4 takes the questionnaire and makes sure they are completely filled out then instructs them to go to a specific work station. Each work station has a desktop computer, a sheet of blank paper, an electronic calculator, an individually packed hand sanitizer wipe, and a bottle of still drinking water.

The laboratory is equipped with a wide-angle HD camera that records the entire session. The footage is used to record the number of times one temporarily removes their masks e.g. to drink water, and the use of hand sanitizer wipe. After the laboratory experiment, participants are paid in cash and called in the order of their ID number, ensuring that they all spend the same amount of time on camera.

Questionnaire Handout

Date and time:

ID number:

Answering all questions is mandatory to complete the experiment. Please keep this sheet and hand it over to the experimenter when you enter the laboratory.

Were you contacted by email and asked to wear a higher-grade mask (N95, FFP2, KF94, KN95, or higher) during this experiment?

☐ Yes

☐ No

If you answered yes to the previous question, did you

☐ bring your own mask

☐ or receive a free mask at the registration?

What sort of mask are you wearing now?

☐ Surgical mask.

☐ Fabric mask (e.g. textile)

☐ Respirator (N95, FFP2, KF94, KN94, or higher).

Instructions

Welcome to the experiment. Please read these instructions carefully. They are identical for all the participants in this session. If you have any questions, please raise your hand, and an experimenter will come to you and answer your questions. Communication with other participants is not allowed. You must leave the experiment with no payment if you do not conform to these rules.

This is an economic experiment. Therefore all the information we give you will be true. There will be no debriefing after the session. During the experiment, you can earn money which will be paid to you in private and in cash at the end of the session. You will receive $5 just for completing the experiment, but you can earn more depending on your decisions and performance in the tasks. The experiment consists of two parts. You must complete the two parts and a questionnaire to receive your earnings.

Part 1

In this part of the experiment, you receive an endowment of $17.50. You might keep this amount until the end of Part 1 or lose $15. Whether you make a loss depends on your decisions and chance. Below we explain how to make decisions and how your payment is calculated.

You make five decisions. With equal chance, only one of these decisions will influence your payment. For example, in a round, you face a decision to buy insurance. You do this by choosing your **insurance number**. A higher insurance number reduces the chance of losing $15, but it is costly. Both the chance of losing money and the cost of insurance depends on a *coefficient* that changes in each round, but they are calculated as follows:

- The cost of insurance is

$$\frac{INSURANCE^2}{COEFFICIENT}$$

- And the chance of losing $15 is

$$\frac{1}{INSURANCE \times COEFFICIENT}$$

Example: If the coefficient is 4 and you choose to buy 1.2 insurance, then you pay $\frac{1.2^2}{4} = 0.36\$$ for the insurance and the chance of losing \$15 will be $\frac{1}{1.2*4} = \frac{1}{4.8} = 20.8\%$. The computer randomly decides if you lose the amount. In this example, you earn 17.50-0.36-15=\$2.14 with a 20.8% chance and 17.50-0.36=\$17.14 with a 79.2% chance.

The coefficient is a number between 4 and 10. Each of the five rounds has a coefficient of unknown value, a different value in all five periods. We do not tell you its value but you can learn about it as follows.

Before making a choice, you can run a trial simulation by producing draws using the same coefficient. On the first screen, you are asked to choose an insurance number. This is a simulation, that is, this choice does not influence your payment. Then, the computer makes ten draws and tells you the outcome: whether you made a loss, the outcome of the draw and your simulated payment in each case.

Only on the next screen, can you make your choice that influences your payment. You can make any permissible choice; the simulated choice you made does not limit you. Note that all draws are independent of each other, but the coefficient is the same in the actual and the ten simulated draws. After this, a new round starts with a new coefficient.

There are two limits: The insurance number you choose must be at least 1 and the amount you spend on the insurance cannot be larger than \$2.50. Hence, you can never lose money in this experiment. You can choose your number up to one decimal precision. This means that you can choose any insurance number between 1 and 3.1.

Remember: Although you make five decisions, only one of them is paid out. You do not know which one, in advance. After the five decisions, you will learn the outcome of this part.

Part 2

[Only read this after Part 1!]

In Part 2, you make 10 decisions. you can see these 10 decision problems on your screen. You can choose between two alternative options for each of the 10 decision problems. Your decision is final only after you have selected an option in each row and have clicked on the “OK” button at the bottom of the screen. Take your time to make your decisions because your choice – as described below – determines your payoff in Part 2.

After you have made all ten decisions and pressed ’ok’, your payoff is determined as follows: The computer will draw two random numbers with a virtual 10-sided dice between 1 and 10, with all numbers equally likely. The first random number determines the row number from the table on your screen. The option you select in this row will be executed using the second random number. You get the outcome from this option at the end of the experiment.

Example: Suppose the computer has selected 2 as the first number, i.e., the decision problem in the second row of the table. Then, the option you have chosen in the second row will be executed, i.e., it will be relevant for your payoff. Suppose you have selected option A in that row. Then, in this case, you will receive either $2 (with 20% probability or if the second random number is 1 or 2) or $1.60 (with 80% probability or if the second random number is 3;4;5;6;7;8;9 or 10). Assume the computer has selected 9 as the second number. Your payoff in Part 2 would thus be $1.60.

You make your decisions only once. The random numbers are drawn at the end of this part. You will learn the outcome.

Questionnaire At the end of the experiment, we will ask you to complete a computerized questionnaire. After completing the experiment and the questionnaire, you will receive your earnings in cash, in addition to the $5 show-up fee.

Supplementary Materials B: Summary Statistics

Table A1: Summary Statistics

	Mean	Standard Dev.	Minimum	Maximum
Safety Interventions				
HEPA	0.51	0.50	0	1
Mask	0.52	0.50	0	1
Outcome Variables				
Distance (cm)	62.37	25.95	13	165
Mask Off	0.86	1.43	0	12
Sanitizer	0.11	0.31	0	1
Characteristics				
CRRA coefficient r_i	0.52	0.027	-0.95	1.37
Gender (Female=1)	0.54	0.50	0	1
Age in Years	23.23	3.57	19	46
Study Master	0.06	0.006	0	1
Study Ph.D.	0.05	0.005	0	1
Questionnaire				
Importance Indoor Masking	6.16	1.09	1	7
Importance Regular Testing	5.66	1.43	1	7
Importance Hot Water	2.11	1.43	1	7
Importance Ventillation	5.82	1.18	1	7
Importance Crowd Control	5.95	1.21	1	7
Importance Air Filtering	5.32	1.35	1	7
Importance Outdoor Masking	4.12	1.85	1	7
Importance Vitamin C	3.41	1.72	1	7
Importance Contact Tracing	4.88	1.64	1	7
Comfort Chair	6.14	1.06	2	7
Comfort Mask	5.48	1.50	1	7
Comfort Temperature	5.96	1.19	2	7

Notes: One half of the sessions received HEPA intervention, in which subjects were told that a high-efficiency particulate absorbing (HEPA) filter was operating. One half of the subjects received MASK intervention where they were required to wear an N95 or equivalent respirator throughout the session. DISTANCE is the distance between the subject and the person in front of them in a queue, before entering the laboratory. MASKOFF is the number of times the subject removed their mask during the experimental session. SANITIZER is a dummy variable, equals 1 if the subject used or took the sanitizer wipe, and 0 otherwise. The questionnaire included questions eliciting the subjects' opinions on a 7-point Likert scale.

Supplementary Materials C: Robustness Checks

Table A2: Estimates on the amount of insurance purchased

	(1)	(2)	(3)	(4)
	$e_{i,t}$	$e_{i,t}$	$e_{i,t}$	$e_{i,t}$
$w_{i,t}$	-0.034***	-0.035***	-0.064**	-0.066**
	(0.0055)	(0.0055)	(0.020)	(0.021)
t		0.025**	0.037**	0.038**
		(0.0083)	(0.014)	(0.014)
$w_{i,t-1}$			0.014	0.017
			(0.020)	(0.021)
r_i				0.049
				(0.069)
Age				-0.074
				(0.075)
Female				0.0070
				(0.070)
High Math				-0.074
				(0.069)
Constant	2.23***	2.17***	2.22***	2.23***
	(0.047)	(0.052)	(0.072)	(0.095)
Observations	1570	1570	1256	1256
R^2	0.011	0.014	0.025	0.040

Notes: Random-effect GLS regressions. Standard errors in parentheses: † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The dependent variable $e_{i,t}$ is the amount of insurance purchased by subject i in period t. The main independent variable of interest is $w_{i,t}$, the safety condition. The CRRA coefficient r_i is estimated with the Holt and Laury (2002) choice experiment. High Math is a dummy variable whose value is 1 for those who took at least 3 university-level math courses. Age is a dummy variable whose value is 1 for those aged above the median age of 23.