Skewness Preferences: Evidence from Online Poker*

Markus Dertwinkel-Kalt
University of Münster &
Max Planck Institute for Research on Collective Goods

Johannes Kasinger
Goethe University Frankfurt &
Leibniz Institute for Financial Research SAFE

Dmitrij Schneider
Düsseldorf Institute for Competition Economics (DICE),
Heinrich Heine University Düsseldorf

October 29, 2022

Abstract

We test for skewness preferences in a large set of observational panel data on online poker games (n=4,450,585). Each observation refers to a choice between a safe option and a binary risk of winning or losing the game. Our setting offers a real-world choice situation with substantial incentives where probability distributions are simple, transparent, and known to the individuals. Individuals reveal a strong and robust preference for idiosyncratic skewness, which has important implications for asset pricing. The effect of skewness is most pronounced among experienced and losing players but remains highly significant for winning players, in contrast to the variance effect.

JEL-Classification: D01, D81, G40.

Keywords: Risk Preferences; Choice under Risk; Idiosyncratic skewness; Behavioral asset pricing.

*We thank Björn Bartling, Sabine Bernard, Thomas Dohmen, Markus Eyting, Katrin Gödker, Andreas Hackethal, Matthias Heinz, Simon Heß, Michael Kosfeld, Mats Köster, Jan Krahnen, Christine Laudenbach, and Matthias Sutter for constructive comments and suggestions. We further thank seminar participants at the Goethe University Frankfurt, the Heinrich Heine University Düsseldorf, and the Leibniz Institute for Financial Research SAFE as well as several conference audiences for valuable feedback. Dertwinkel-Kalt: University of Münster, markus.dertwinkel-kalt@wiwi.uni-muenster.de. Kasinger: Goethe University Frankfurt and Leibniz Institute for Financial Research SAFE, kasinger@safe-frankfurt.de. Schneider: Düsseldorf Institute for Competition Economics (DICE), Heinrich Heine University Düsseldorf, schneider@dice.hhu.de.

1 Introduction

According to standard portfolio theory (Markowitz, 1952), mean-variance preferences reflect the main trade-off for choice under risk. More recent literature in finance and behavioral economics has argued, however, that also higher moments play a decisive role—in particular skewness, a distribution's standardized third moment. There is compelling empirical evidence for *skewness preferences*—preferences for positive and an aversion toward negative payoff skewness (Boyer *et al.*, 2010; Bali *et al.*, 2011; Green and Hwang, 2012; Conrad *et al.*, 2013; Boyer and Vorkink, 2014; Lin and Liu, 2018; Jondeau *et al.*, 2019). These empirical findings challenge traditional asset pricing models, in which only coskewness with the market is relevant for asset prices (Kraus and Litzenberger, 1976; Harvey and Siddique, 2000). In contrast, more recent behavioral-based asset pricing models depart from the standard expected utility framework and posit that an asset's idiosyncratic skewness is priced because investors have a preference for skewness or lottery-like features of assets (for such models see Brunnermeier and Parker, 2005, Mitton and Vorkink, 2007, and Barberis and Huang, 2008). Despite the vast empirical literature on skewness preferences, the relevance of idiosyncratic skewness for financial decision-making and asset prices remains in question as alternative explanations of the identified effects can usually not be ruled out. Measuring ex-ante skewness in real-world settings is usually impossible, and risk preference estimation requires strong assumptions because the underlying probability distribution is usually very complex, estimated based on past data, and de facto unknown to the decision-maker.

This paper directly tests for skewness preferences in a large set of observational panel data on real-world choices. The studied decision can be understood as a stylized example of an investor's decision between a risk-free investment and an investment in an idiosyncratic "risky asset". The decision situations we analyze are subject to a well-defined set of lotteries, whose underlying ex-ante probability distribution can be objectively and unambiguously calculated. In addition, probabilities are transparently displayed to decision-makers at decision time.[1] Consequently, we do not need to approximate the underlying probability distribution and impose strong assumptions that individuals correctly estimate and comprehend such a distribution.

[1] Our setting shares many desirable features with betting markets (e.g. Andrikogiannopoulou and Papakonstantinou, 2020; Moskowitz, 2021). For example, termination points and terminal payoffs of bets are independent of investors' beliefs, preferences, or other factors, such as systematic risk. One major difference in betting markets is that ex-ante objective probabilities cannot be observed.

The risks involved are binary and uniquely determined by the first three moments of their probability distribution: expected value, variance, and skewness, allowing for a clean identification of skewness preferences. In addition, individuals in our data set can generally be expected to understand risks and probabilities well. Thus, we regard our data set as well-suited to study risk *preferences* (rather than risk-taking driven by misunderstandings or imperfect information).

We show that skewness has a sizeable and significant effect on risk-taking, while the effect of variance is comparably negligible. These findings are not coherent with expected utility theory (see Section 2) and lend support to behavioral asset pricing models that allow for preferences for skewness or lottery-like features of stocks. Skewness prefernces are a central prediction of most behavioral models of choice under risk, such as prospect theory (Kahneman and Tversky, 1979), disappointment aversion (Gul, 1991), or salience theory (Bordalo *et al.*, 2012). They can explain a wide range of seemingly disparate puzzles in finance, such as the growth puzzle (Fama and French, 1992; Bordalo *et al.*, 2013), many instances of portfolio underdiversification (Mitton and Vorkink, 2007), and the disposition effect (Barberis and Xiong, 2009; Fischbacher *et al.*, 2017).[2] Moreover, there are numerous behavioral studies providing evidence for skewness preferences in stylized laboratory experiments with limited generalizability to real-world economic behavior (e.g., Ebert and Wiesen, 2011; Ebert, 2015; Trautmann and van de Kuilen, 2018; Dertwinkel-Kalt and Köster, 2020). Unlike laboratory experiments on skewness preferences, we cover a wider range of incentives, have a higher number of observations, and do not build on an artificial choice situation.

In detail, we study risk-taking in online poker, making use of a novelty that the world's largest online poker platform, *Pokerstars*, introduced in August 2019: the so-called *all-in cashout*. The all-in cashout provides insurance against a player's risk in a showdown situation. In a showdown situation, the outcome of the poker hand is solely determined by the cards drawn from the remaining deck of cards. All relevant information is disclosed, and the respective winning probabilities for each player can be easily calculated. The two outcomes of the implied binary lottery are: i) losing and receiving a payout of zero or ii) winning the entire *pot*, i.e., the accumulated bets by players at the table throughout a hand. The all-in cashout gives each

[2] Other examples of puzzles in choice under risk that can be explained by skewness preferences include the favorite-longshot bias, whereby positively skewed long shots are overbet and negatively skewed return distributions of favorites are underbet (e.g., Snowberg and Wolfers, 2010), the simultaneous demand for lottery-like gambles and high-premium insurances (Kahneman and Tversky, 1979; Sydnor, 2010; Garrett and Sobel, 1999) and the Allais paradox (Allais, 1953). Test of skewness preferences in labor economics include (Hartog and Vijverberg, 2007; Berkhout *et al.*, 2010).

player in a showdown situation the option to choose a safe payout equal to the expected payout of the underlying lottery minus a profit margin for *Pokerstars* of 1%. Each observation refers to a player's choice between this safe insurance option and the respective binary lottery.[3] Our data set includes 4,450,585 of such individual choices in showdown situations with two opponents for Omaha Poker cash games, all collected between January 1, 2020, and June 30, 2021.[4]

As winning the pot is a complementary event for the two opponents, for each observation involving a right-skewed lottery with the winning probability $\pi < 0.5$, there is exactly one observation in our data set that involves a left-skewed lottery with the winning probability $1 - \pi > 0.5$. Both of these lotteries have identical variance but inverse skewness, which allows us to circumvent the pitfalls of other studies (e.g., Golec and Tamarkin, 1998) where any choice shift that is attributed to a change in skewness could be likewise attributed to a change in variance.[5] Unlike variance, skewness has a strong and robust effect on risk-taking. The insurance option is selected in 20.0% of cases when the risk is left-skewed, but only in 14.2% of cases when the risk is right-skewed. Put differently, it is around 40% more likely that the insurance option is selected when players face a left-skewed instead of a right-skewed risk. This difference is highly statistically significant (p-value < 0.0001, t-test).

In our regression analyses, we follow Mitton and Vorkink (2007) and assume a utility function that is linear in the different risk moments. In our basic specification, we regress an insurance choice dummy that equals one if player i chooses the insurance option in showdown j and zero otherwise, on the first three moments of the underlying lottery. We further include game fixed effects to control for heterogeneity in insurance choices across different games and month fixed effects to account for month-specific heterogeneity as potentially driven by seasonal effects or COVID-19 countermeasures. Increasing skewness by one standard deviation, keeping the variance and expected values of the lotteries equal, decreases the likelihood that the insurance option is chosen by 2.3−2.4 percentage points, which is equivalent to a decrease of around 13.5% compared to the average share of positive insurance choices (i.e., the mean dependent

[3]Strictly speaking, the lottery is not always binary as there are situations where a split pot can occur; we discuss this limitation of our study in Section 3.3.

[4]Showdowns with more than two players typically yield more complex probability distributions, depending on when each player decides to go all-in and depending on each player's budget when going all-in. Thus, for clarity, we restrict our analysis to situations where only two players go all-in.

[5]In our study, in every single choice situation, the right-skewed lottery always has a smaller expected value than its left-skewed counterpart, making it impossible to distinguish between the effect of a larger skewness versus the effect of a lower mean in a given choice situation. While we account for different expected values in our regression analyses, neither standard nor behavioral models suggest that the mean should be an important driver of risk-taking in our setup.

variable). The estimated effect of variance is negligible in most specifications. A one standard deviation increase in variance is associated with a decrease in insurance choice of around 0.0 to 0.3 percentage points, depending on the specification. In the basic linear probability model, the estimated effect is only weakly significant. These results remain robust to different empirical specifications (such as Probit and Logit models), to controlling for player-specific characteristics (such as experience and average profit per hand) and hand-specific variables (such as the amount of money the player started the hand with, the weekday, or the *stake*, i.e., the size of mandatory bets), to excluding outliers, and to using the coefficient of variation (the inverse "Sharpe" ratio of the lottery) instead of expected value and variance. The panel structure of our data further allows us to include individual fixed effects to control for time-invariant heterogeneity across individuals. Including individual fixed effects alter the effect of skewness, neither in terms of magnitude nor significance, and does not affect the estimated coefficient of the first two moments by much.

To learn more about the robustness, the generalizability, and the implications of our results, we study heterogeneity in risk preferences in different subgroups. We split our sample at the median for the various player- and hand-specific characteristics. We find evidence of skewness preferences in all subsamples. The evidence is strongest for relatively experienced players, as measured by the number of showdowns the player participated in or by the total number of hands played (including hands without showdowns). For this subsample, increasing skewness by one standard deviation decreases the likelihood that the insurance option is chosen by around 3.3 percentage points, which is about one and a half times the effect estimated in the full sample. Moreover, when restricting the analysis to the experienced players, the effect of the other two moments decreases in absolute size or even changes signs. Related studies (Palomäki *et al.*, 2013, 2014) suggest that experienced poker players are more self-reflective, less affected by emotions, and likely to understand the decision environment and the underlying probabilities better than inexperienced players. Taking these insights into account, we think that the estimated effect of a moment for an experienced player can be plausibly attributed to a *preference* for that moment, which is rather not diluted by a misperception of the underlying lotteries or by random and emotional choice.[6]

[6] For this reason, Eil and Lien (2014) focus exclusively on experienced poker players in their study on reference-dependent risk attitudes.

While a player's skill cannot affect the outcome of the studied lotteries, it can affect the long-run profits of players. To investigate how far our findings depend on a player's skill, we run our analyses separately for winning and losing players. Winning (losing) players have made a positive (negative) net profit over our observation period playing Omaha Poker cash games. To address the issue that poker profits can be the results of luck in the short run, we further define winning players as sophisticated if, in addition to making positive profits, they are experienced (i.e., they have played sufficiently many hands). Likewise, we define recreational players as experienced but losing players.[7] Only the effect of skewness is significant in all subsamples, but more pronounced for recreational players.

Despite the popularity of online poker, our data set relies on a selective sample of people—a feature that is shared by most field studies on risky choice, which are restricted to, e.g., financial investors (e.g. Boyer *et al.*, 2010; Conrad *et al.*, 2013; Lin and Liu, 2018), game show participants (e.g. Gertner, 1993; Post *et al.*, 2008), bettors (e.g. Snowberg and Wolfers, 2010; Andrikogiannopoulou and Papakonstantinou, 2020), or people buying auto insurance (e.g. Cohen and Einav, 2007). In our case, the selected sample has advantages and disadvantages. On the upside, poker players are used to risky choice situations and probabilities. Thus, our findings should not be confounded by a misunderstanding of the lotteries and can be plausibly attributed to the player's risk preferences. On the downside, online poker may attract individuals with non-representative (risk) preferences, which is backed by the observation that the overall insurance take-up is rather low and playing online poker has, on average, a negative expected return due to the fee taken by the platform providers (on average a player loses around \$0.088 per hand). Past studies suggest that (online) poker players are more likely to be younger and male than the general population (Barrault and Varescon, 2016). While not being necessarily applicable to the general population, our results could be rather informative for individuals that self-select into risky choices in other instances, such as individual investors—particularly those with a large propensity to invest in lottery-type stocks (Kumar, 2009) or cryptocurrencies (Hackethal *et al.*, 2022)—, bettors (e.g. Andrikogiannopoulou and Papakonstantinou, 2020; Moskowitz, 2021), and entrepreneurs (Moskowitz and Vissing-Jørgensen, 2002).[8]

[7]Sophisticated and recreational players may systematically differ in risk attitudes and preferences. sophisticated players may resemble financial experts, while recreational players may resemble speculative retail investors better.

[8]That our results carry over to other relevant populations is also supported by the fact that similar findings have also been documented in an experimental laboratory study with German University students (see Experiment 1 in Dertwinkel-Kalt and Köster, 2020).

In addition to the broad literature on skewness preferences in finance, behavioral economics and labor economics, this paper also adds to the literature on decision making in poker games. Besides some studies that aim to quantify the extent of skill and luck in poker games (e.g., Fiedler and Rock, 2009; Potter van Loon *et al.*, 2015; Duersch *et al.*, 2020), researchers have mainly used poker data to study reference-dependent risk attitudes. Smith *et al.* (2009) and Eil and Lien (2014) find that poker players play less cautiously, longer, and more aggressively after losing a big pot or if a player is losing within a poker session. These findings are in line with the break-even hypothesis (Kahneman and Tversky, 1979; Thaler and Johnson, 1990). One challenge when analyzing poker play is that it crucially depends on the players' expectations about the opponents' hands and playing style (both of which are unobservable). By focusing on those situations where all uncertainty is resolved and players cannot actively affect the outcome of the game anymore, we can, however, overcome these problems. So unlike in previous studies (e.g. Smith *et al.*, 2009; Eil and Lien, 2014), missing information, wrong beliefs or a misunderstanding of the risks involved should not play a role in our setup and should not confound our insights on the drivers of risk-taking.

We proceed as follows. In Section 2 we define skewness preferences and discuss the related literature. In Section 3 we describe our data before we present our results in Section 4. Section 5 concludes.

2 Theoretical Background and Related Literature

Our observations involve binary decisions between a safe "insurance" option and a binary lotteries L. As we will see in Lemma 1, such a binary lottery is uniquely defined by its expected value $\mathbb{E}[L]$, its variance $Var[L]$ and its skewness $S[L]$, which is defined by the third standardized central moment

$$S[L] := \mathbb{E}\left[\left(\frac{L - \mathbb{E}[L]}{\sqrt{Var[L]}}\right)^3\right]. \tag{1}$$

We can then define the following notions.

Definition 1. *Lottery L is called* right-skewed *(or, equivalently, positively skewed) if $S(L) > 0$,* left-skewed *(or, equivalently, negatively skewed) if $S(L) < 0$, and* symmetric *otherwise.*

Other definitions of positive skewness, such as via "long and lean" tails of the risk's probability distribution, exist and are in general not equivalent. For binary risks $L = (x_1, \pi; x_2, 1-\pi)$, where outcome x_1 is realized with probability π and x_2 with probability $1 - \pi$, Ebert (2015), however, shows that all conventional notions of skewness are equivalent and the skewness of a binary risk is well-defined. Moreover, binary lotteries can be uniquely identified by their first three moments as shown by Ebert (2015) and generalized by Dertwinkel-Kalt *et al.* (2021):

Lemma 1. *For constants $E \in \mathbb{R}$, $V \in \mathbb{R}_+$ and $S \in \mathbb{R}$, there exists exactly one binary lottery $L = (x_1, \pi; x_2, 1-\pi)$ with $x_2 > x_1$ such that $\mathbb{E}(L) = E$, $Var(L) = V$ and $S(L) = S$. Its parameters are given by*

$$x_1 = E - \sqrt{\frac{V(1-\pi)}{\pi}},\ x_2 = E + \sqrt{\frac{V\pi}{1-\pi}},\ \text{and } \pi = \frac{1}{2} + \frac{S}{2\sqrt{4+S^2}}. \tag{2}$$

We denote the binary lottery with expected value E, variance V, and skewness S as $L(E, V, S)$.

As a result of Lemma 1, it is possible to vary skewness for binary risks while fixing the first two moments. We can now define skewness preferences as follows:

Definition 2. *For any $E \in \mathbb{R}$ and $V \in \mathbb{R}_+$, an agent reveals a preference for skewness if the following holds: she strictly prefers the binary lottery $L(E, V, S)$ over the safe option that pays the lottery's expected value E if and only if S is strictly positive and sufficiently large.*

Expected Utility Theory. In the following, we formally show that Definition 2 violates expected utility theory (henceforth: EUT) and is not coherent with *any* EUT utility function. Let $u(\cdot)$ be a utility function with normalization $u(0) = 0$, and consider $L = (0, 1-\pi; x, \pi)$. According to Definition 2, there exists a threshold skewness value, or equivalently, a threshold probability level $\bar{\pi} \in \left(\frac{1}{2}, 1\right)$ so that for all $x \geq 0$ and all $\pi \in (0, 1)$ we have

$$u(\pi x) \geq \pi u(x) \text{ if } \pi > \bar{\pi} \tag{3}$$

$$u(\pi x) < \pi u(x) \text{ if } \pi \leq \bar{\pi}. \tag{4}$$

Now we define $x' := 2x$ and assume $\pi \in (\frac{1}{2}, \frac{\bar{\pi}}{2})$. Then, (3) and (4) yield $2\pi u(x) \leq u(2\pi x) = u(\pi x') < \pi u(2x)$, and thus $2u(x) < u(2x)$. But we also have $u(x) = u(\frac{1}{2} \cdot 2 \cdot x) \geq \frac{1}{2} u(2x)$, which gives $2u(x) \geq u(2x)$. Taken together, these two derivations give a contradiction. Intuitively, Definition 2 stipulates that the agent is (i) risk-averse over all risks that are not sufficiently skewed, which necessitates a weakly concave utility function, but (ii) risk-seeking over all sufficiently skewed risks, which necessitates a strictly concave utility function. Taken together, (i) and (ii) give a contradiction.

Notably, in our empirical setting the agent cannot obtain a gamble's expected value E, but only slightly less, namely $0.99E$. This, however, does not alter any of the conclusions we have drawn on the validity of EUT, as demonstrated in the following. Suppose an agent strictly prefers a risky option $L = (0, 1 - \pi; x, \pi)$ over 99% of its expected value, πx, if and only if the risky option's skewness is sufficiently large. Then we obtain the conditions that are analogous to (3) and (4), namely,

$$u(0.99\pi x) \geq \pi u(x) \text{ if } \pi > \bar{\pi}$$

$$u(0.99\pi x) < \pi u(x) \text{ if } \pi \leq \bar{\pi}.$$

The very same contradiction as above can be constructed from these conditions. Intuitively, in order to weakly prefer $0.99E$ over a risky option with expected value E and non-positive skewness, the utility function must be sufficiently concave, but in order to prefer the risky option over $0.99E$ when skewness is sufficiently positive, the utility function must not be that concave—a contradiction.

EUT with the only additional assumptions of prudence (whereby the third derivative of the utility function is strictly positive, that is, $u''' > 0$) can, however, explain why agents strictly prefer, c.p., right- over left-skewed risks, so $L(E, V, S)$ over $L(E, V, -S)$. Some papers in the literature, such as Ebert and Karehnke (2020), use this definition of skewness preferences, whereby—when choosing between two binary lotteries $L(E, V, S)$ and $L(E, V, -S)$ that are identical in all but the sign of skewness—the lottery $L(E, V, S)$ is chosen. Unlike the definition that we use, skewness preferences in this alternative sense can be reconciled with EUT (by assuming the right signs for higher order derivatives of the utility function; see Kraus and Litzenberger (1976) for a classical finance model that presents this approach).[9] We, however, view this alternative definition as less relevant for practice, as our motivating examples (gambling vs. insurance choice or most of the applications in financial or labor economics) as well as the choice situations that we investigate better fit the definition we build on. We are not even aware of any

[9]For financial markets, Kraus and Litzenberger (1976) can explain within EUT why a tradable financial asset's coskewness with the market is priced, but not why its own skewness is priced. To explain this, behavioral economic models are needed as pointed out, for instance, by Barberis and Huang (2008).

real-life situations where such choices between $L(E, V, S)$ and $L(E, V, -S)$ play a role, that is, where an agent chooses between two risks that just differ in the sign of skewness; all common examples (whether to gamble or to buy insurance, for instance) instead reflect the setup of Definition 1, that is, the choice between *one risky* and *one safe* option.

Taken together, skewness preferences as defined in Definition 2 are not coherent with EUT. In order to explain why agents dislike symmetric risks, EUT needs to assume that the utility function is strictly concave. Under this assumption, however, EUT predicts, for any skewness level, that the safe payout of E should be strictly preferred over any lottery that pays E in expectation. When EUT wants to explain why an agent takes up a positively skewed risk, it has to assume a convex utility function ($u'' > 0$), but this then goes along with the implausible prediction (that also violates Definition 2) that *every* binary lottery is preferred over the safe option that pays its expected value. So, EUT cannot explain why people's preference to take up a risk depends on the risk's skewness. The same holds for standard portfolio theory (Markowitz, 1952), whereby people's utility from some risk is linearly increasing in its expected value and linearly decreasing in its variance. Thus, EUT and standard portfolio theory would predict that whether some risk is taken up mainly depends on variance, but not on skewness.

Behavioral Economics. Skewness preferences as defined in Definition 2 are predicted, however, by most behavioral models of choice under risk such as cumulative prospect theory (Kahneman and Tversky, 1979), salience theory (Bordalo *et al.*, 2012), regret theory (e.g. Bell, 1982; Loomes and Sugden, 1982), and disappointment aversion (Gul, 1991), as shown, for instance, in Barberis (2012), Dertwinkel-Kalt and Köster (2020), and Ebert and Karehnke (2020). Also seminal models proposed in the behavioral finance literature (Mitton and Vorkink, 2007) allow for skewness preferences by, for instance, augmenting standard portfolio theory by an additional term that allows not only for a linear effect of expected value and variance, but also of skewness on utility.

Skewness preferences allow us to understand why revealed attitudes toward risks vary across contexts. On the one hand, people like to gamble (e.g., Golec and Tamarkin, 1998, Garrett and Sobel, 1999) but they also overpay for insurance with low deductibles (e.g., Sydnor, 2010; Barseghyan *et al.*, 2013). In financial markets, investors seek positively skewed return distributions (Chunhachinda *et al.*, 1997; Prakash *et al.*, 2003; Mitton and Vorkink, 2007; Boyer *et al.*, 2010; Bali *et al.*, 2011; Conrad *et al.*, 2013). Skewness preferences also matter in labor economics as they affect career choices (Hartog and Vijverberg, 2007; Berkhout *et al.*, 2010; Grove *et al.*, 2021) as, for instance, workers accept a lower expected wage if the distribution of wages in a cluster (i.e., education-occupation combination) is right-skewed. Similarly, a preference for skewness can explain the substantial entrepreneurial investments in private equity with bad risk-return tradeoffs (Moskowitz and Vissing-Jørgensen, 2002).

The characteristics of binary lotteries outlined above make it appealing to study skewness effects at the hand of binary lotteries. To identify skewness preferences, it is optimal to let agents repeatedly choose between a safe option and a binary lottery, where only the lottery's skewness differs between choices. In such a setting, skewness preferences predict a negative relation between insurance choice and the lottery's skewness. While this decision situation is hardly implementable in the field (for a laboratory experiment that precisely implements this see Experiment 1 in Dertwinkel-Kalt and Köster, 2020), our setup approximates these experiments as closely as possible (for a discussion of differences to the ideal setup see the discussion of limitations in Section 3.3).

In our setup, poker players face the choice between a lottery and the safe option that pays 99% of the lottery's expected value. Thus, the lottery is selected if and only if for the player's utility function $U(\cdot)$ we have $U(L) > 0.99\ U(\mathbb{E}(L))$. We follow Mitton and Vorkink's (2007) reduced-form approach by assuming "Lotto investors" that have identical preferences as traditional investors over mean and variance (see Markowitz, 1952), but also a preference for skewness. The utility such investors derive from some lottery L is then given by

$$U(L) = \mathbb{E}(L) + \beta_V Var(L) + \beta_S S(L).^{10}$$

A positive (negative) coefficient β_V on variance indicates variance-loving (variance-averse) agents, and a positive (negative) coefficient β_S on skewness indicates a preference for positive (negative) skewness. This approach therefore allows for both positive and negative effects of variance and skewness on a lottery's valuation. We will not directly estimate the effect of the lottery's moments on the lottery's utility, but on the likelihood that the lottery is preferred over the safe alternative.

[10] Unlike Mitton and Vorkink (2007) we adopt the usual narrow-framing assumption that is adopted throughout experimental economics: namely, that subjects do not integrate their earning from the respective game into their overall wealth, but evaluate it in isolation.

3 Background and Data

In this section, we first provide the background on *Omaha Poker* by explaining the game and the new insurance option that our study builds on. We then give an overview of our data set and discuss its advantages and limitations.

3.1 Background

On Omaha Poker cash games

We analyze hands from Omaha Poker cash games. In a cash game, all players start the hand with an amount of real money, the *stack*, which will be used for betting throughout the respective hand.[11] Money cannot be added or withdrawn during a hand. However, players can leave the table after a hand is concluded or add chips up to a maximum amount depending on the blinds—size of mandatory bets posted before every hand—of the respective table. Accordingly, we define the stake of a game by the size of the blinds.[12]

In Omaha Poker cash games, each player is dealt four private cards (*hole cards*) that are only visible to the respective player. In addition, there are up to five *community cards* that are public information and are dealt throughout three stages: i) *Flop*: first three community cards; ii) *Turn*: fourth community card; iii) *River*: fifth community card. Each stage is preceded and/or followed by a betting round. The money that players post throughout these betting rounds is collected in the *pot*. Furthermore, there is a fee collected by Pokerstars, called the *rake* that is deducted from the pot. The rake is calculated as a percentage of the pot, ranging from 3.5% to 5% depending on the stake, and capped at a certain amount.[13] Accordingly, the winning player is awarded the *net pot*, that is, the pot minus the rake.

If the betting causes all but one player to lay down their hole cards (i.e., they *fold*), the remaining active player wins the net pot without showing any private cards. Otherwise, the net pot is awarded to the active player with the best five-card poker hand after the last community

[11] Beside the fact that the insurance option is only available for cash games, cash games are also more suited to our question and "easier to analyze than tournament games, since in a cash game, a player who is risk neutral over money should also be risk neutral over chips. This is not necessarily the case in a tournament, for a number of reason" Eil and Lien (2014), including varying incentives to outlast other players in different tournament phases or non-linear payout structures.

[12] In a poker cash game there are usually two blinds, the *big blind* and the *small blind*, which is half the size of the big blind. In the remainder of the paper, the stake always refers to the big blind.

[13] For more details see: *https://www.pokerstars.eu/poker/room/rake/*.

card is dealt. This best five-card poker hand consists of two of the player's hole cards and three community cards (see also the official ranking of poker hands in Appendix A.1).

The hole cards are revealed when either the betting round after the River is finished or when there are $N > 2$ active players, of which at least $N - 1$ players are *all-in*, that is, they put their entire stack in the pot. The latter scenario is called a *showdown*. In a showdown, the players face a binary lottery L, whose outcome depends solely on the cards that will be drawn from the remaining deck of cards: $L = (\text{pot} - \text{rake}, \pi; 0, 1 - \pi)$.[14] In such a situation, the probability that one player wins the net pot, π, can be calculated conditional on the revealed individual hole cards and the community cards that have been dealt until the showdown.

Figure 1: Exemplary screen of a showdown on Pokerstars with one card to come

Figure 1 shows an example of a showdown after the Turn, i.e., with one card to come. Player 1 is all-in, and no more betting is possible. The best five-card hand of Player 1 is a *High Card Queen*, which loses against the *Two Pairs* (queens and tens) of Player 6. Player 1 can only win the hand if a *Heart*-card is drawn from the remaining cards, which would give her a winning *Flush*

[14]Under certain circumstances split pot situations may occur. In such a scenario, the lottery is not binary. We abstract from these scenarios here and discuss the issue in Section 3.3 in more detail.

(five cards of the same suit). In total, there are 13 Heart-cards in the 52-card deck. Five Heart cards have been already revealed, implying that there are still eight Heart cards among the 40 cards that have not been revealed yet. The probability of Player 1 winning the hand is thus simply the number of remaining Heart cards (eight in our example) divided by the number of remaining cards, $52-12 = 40$: $\frac{8}{40} = 0.20$. The probability of losing is accordingly 0.80. As shown in Figure 1, these probabilities and the exact size of the net pot are displayed in an all-in situation on the players' screens. If the showdown happens at an earlier stage, the probabilities can be calculated by dividing all possible realizations of community cards, in which a specific player holds the winning hand, by the total number of possible realizations. Again, the corresponding probabilities and the net pot are displayed to the players (see in Appendix A.2 an example where the showdown happens before any community is revealed).

The insurance option

We make use of the new insurance option (the so-called *all-in-cashout*) introduced on August 13, 2019, on the Pokerstars website, which provides a safe alternative against the risk that the players face in a showdown. If a player chooses the insurance option, she will no longer be eligible to contest any portion of the net pot, and the offered amount will be added to her stack immediately and risk-free. If she declines, she will continue to contest the entire net pot as usual. Players declining the insurance option still need the best hand in a showdown to win the net pot, even if all their opponents have cashed out. As a result, each active player in a showdown faces a choice between a binary risk and a safe option. The guaranteed payout from choosing the insurance option is equal to the expected value of the lottery minus a fee of 1% on this expected value, i.e., equal to $\$(\text{pot} - \text{rake}) \times \pi \times 0.99$. The 1% fee charged by Pokerstars is equal for all players and has not changed since the insurance option's introduction. The insurance payout is rounded to full cents.

To better understand a player's decision in a showdown, turn again to Figure 1, which shows a situation in which the insurance option is offered to Player 1. The two red buttons signify the binary choice between two lotteries: i) the risky option "Resume hand" (explained above) that pays \$2.79 with 20% and zero with 80%; ii) and the safe insurance option that pays \$0.55 with 100%. As can be seen in Figure 1, the probabilities and the exact size of the net pot are displayed on the players' screens. The displayed insurance payout does already include the rake and the

1% fee. Accordingly, Player 1's insurance payout in our example is equal to: $\$2.79 \times 0.2 \times 0.99 = \$0.55$ and Player 6's insurance payout will be: $\$2.79 \times 0.8 \times 0.99 = \$2.21$. Thus, the players are readily presented with all information that is relevant for their decision. The players have 12 seconds to make their choice. If players do not choose one of the proposed alternatives in time, the hand resumes with the risky option.[15]

3.2 Data

Our data set includes 4,450,585 observations, where every observation refers to a unique decision by a single player in a two-person showdown situation as described above. This includes the decisions of 83,219 distinct players.[16]

Our data set is extracted from 35,529,631 distinct poker hands played between January 01, 2020, and June 30, 2021, on Pokerstars, the largest online poker network in the world.[17] We obtain the raw data from a commercial poker data provider that collects and stores hand histories for every Omaha Poker cash game played on Pokerstars.[18] Hand histories are automatically generated by the Pokerstars software and include all public information of a single poker hand.[19] Our data set spans a wide range of tables that vary in the size of the mandatory bets, and the maximum amount players can bring to the respective table. The hand histories include information on all community cards, the hole cards of the players that went to a showdown, and the net pot size. This information allows us to calculate each player's probability of winning the pot at showdown and assign each player's insurance payout value. Furthermore, our data allow us to infer whether a respective player has chosen the insurance option for each showdown.

We capture a player's decision between the safe option and the binary lottery by the variable *insurance choice*, our dependent variable of interest. This variable equals one if the player chooses the safe option and zero if she chooses the binary lottery. In our data set, players choose the safe

[15]For more details see: *https://www.pokerstars.com/poker/promotions/all-in-cash-out*.

[16]As we explain in Section 3.3 in more detail, we exclude all observations that are not skewed, i.e., have a winning probability of 0.5, as most of these observations refer to situations with degenerated underlying lotteries, for which no insurance option was offered.

[17]Note that the number of poker hands exceeds the number of observations as not every poker hand results in a two-person showdown. See the worldwide Online Poker Sites Traffic Report for statistics on the largest online poker platforms: *https://www.pokerscout.com/*.

[18]Our data provider *HH Dealer* (*https://www.hhdealer.com/*) collects the hand histories for various online poker platforms. At Pokerstars, a hand history is accessible by all Pokerstars users that opened the window of a respective table. As a quality check, we played several sessions on the Pokerstars platform during the data collection period and checked whether the obtained histories are complete and accurate.

[19]The raw data is provided in text files and we make use of the commercial poker software "PokerTracker 4" (*https://www.pokertracker.com*) to convert the raw data into a workable data set. Smith *et al.* (2009) use a earlier version of "PokerTracker" to construct their data set.

option in 17.1% of cases with a standard deviation of 0.377 (Table 1). Notably, the overall share of choices of the insurance option is rather low. A majority of players choose the risk instead of the insurance option, which might be driven by a large share of individuals that are risk-seeking in general or by the 1%-margin that has to be paid to *Pokerstars* if the insurance option is chosen.

Table 1: Summary statistics on insurance choice

Statistic	N	#(Choice=1)	#(Choice=0)	Mean	St. Dev.
Insurance Choice	4,450,585	761,585	3,689,000	0.171	0.377

Note: The table reports summary statistics on the insurance choice dummy that equals one if the safe option is chosen and zero otherwise.

The winning probabilities and the net pot size allow us to calculate the expected value, $E = \pi x$, the variance, $V = \pi(1-\pi)x^2$, and the skewness, $S = \frac{1-2\pi}{\sqrt{\pi(1-\pi)}}$, of each binary gamble players face in a showdown situation. Table 2 presents the descriptive statistics for the first three moments of the binary lotteries in our data set. Our lotteries have an average expected value of 62.39, with a standard deviation of 326.58 and a median of 13.37 (all in US-$). The range of the expected values of the different lotteries is very wide and stretches from $0.001 to $64,242.40. Our measure of skewness takes values from -40.79 to 40.79 with a zero mean.

Table 2: Summary statistics on different moments

Statistic	N	Mean	St. Dev.	Pctl(25)	Median	Pctl(75)
Expected Value	4,450,585	62.39	326.58	5.11	13.37	35.92
Variance	4,450,585	73,786.68	2,157,525.00	29.52	148.86	964.71
Skewness	4,450,585	0.00	2.23	−0.86	0.00	0.86

Note: The table reports summary statistics of the expected value, variance, and skewness for the lotteries in our sample. In calculating we use the values of net pot sizes, which are measured in US-$ terms.

The data set also includes each player's screen name, which serves as a unique identifier for each observation (together with the distinct hand number). It allows us to control for individual fixed effects. In addition, our data records each player's stack at the beginning of the hand and some hand-specific characteristics, such as date, time, and the size of the mandatory bets (stake). For all four betting rounds, we further observe the actions of all active players, i.e., players that have not folded their hands. On average, a player in our sample plays 2,159 hands and faces 53.48 two-person showdown situations during the respective period. The average stack at the beginning of the hand is $115.1, ranging from $0.10 to $81,644.[20]

For additional analyses and robustness checks, we extract several player-specific characteristics from the raw data, including the number of hands played over the observation period, the amount won/loss over the entire period (including hands with no showdown), and the average winning probability in showdown situations. As a result of the rake collected by Pokerstars, a player makes an average loss of $0.088 per hand played. More details and the respective summary statistics are provided in Appendix A.4.

3.3 Limitations of our framework

Poker players are not representative of the general population. A player in our sample loses on average $0.088 per hand, and the general likelihood of choosing the insurance option is rather low. Previous studies suggest that (online) poker players are quite likely "young men, executives or students, mostly single and working full-time" (Barrault and Varescon, 2016). Alternative field studies, however, usually also build on non-representative samples.[21] Consequently, we expect our sample to be more representative of people that self-select into related risky choices such as individual investing—in particular in lottery-type stocks (Kumar, 2009) and cryptocurrencies (Hackethal *et al.*, 2022)—, betting (Andrikogiannopoulou and Papakonstantinou, 2020; Moskowitz, 2021), and entrepreneurial investments (Moskowitz and Vissing-Jørgensen, 2002; Vereshchagina and Hopenhayn, 2009). To address generalizability concerns, we also conduct heterogeneity analyses in Section 4.4. Finally, we are confident that our results do not crucially rely on the selection of our sample, as our core results are also detected in (small-scale) lab experiments with German University students (see Experiment 1 in Dertwinkel-Kalt and Köster, 2020, and replications thereof).

Some features of our selection are even advantageous: poker players are generally experienced in dealing with risks and probabilities, so a misunderstanding of the involved lotteries should not confound our identification of moment preferences.

[20] The average stack and the average number of showdown situations are calculated using the 4,450,585 observations in our final data set. The number of hands played by a distinct player is based on all hands in our initial data set, including hands that did not result in a two-person showdown with a winning probability $\neq 0.5$.

[21] Examples include studies investigating financial investors (e.g. Boyer *et al.*, 2010; Conrad *et al.*, 2013; Lin and Liu, 2018), game show participants (e.g. Gertner, 1993; Post *et al.*, 2008), bettors (e.g. Snowberg and Wolfers, 2010; Andrikogiannopoulou and Papakonstantinou, 2020; Moskowitz, 2021), or people buying auto insurance (e.g. Cohen and Einav, 2007).

One limitation of the setup that we investigate is that the underlying risks are, strictly speaking, not always binary as split pots can be possible. Split pots arise when players hold the same best five-card hand after all community cards are dealt. In this case, each involved player that ended up in a showdown is awarded half of the pot. In our sample, 6% of all showdown situations result in a split pot.[22] In these scenarios, players face the following trinary lottery: $L = (x, \pi; \frac{1}{2}x, \mu; 0, 1-\mu-\pi)$, where μ is the ex-ante probability of a split pot and π is the ex-ante probability of winning the entire pot. In our data, μ and π are not independently observable, neither for us nor for the players. In fact, we only observe a "payout-weighted" winning probability $\tilde{\pi} = \frac{1}{2}\mu + \pi$, which can be understood as the percentage of the pot the player is expected to win. If a split pot is possible, the agent thus sees the binary lottery: $\tilde{L} = (x, \tilde{\pi}; 0, 1-\tilde{\pi})$, for which $E(L) = E(\tilde{L})$. In Appendix A.3 we present an example of a choice situation where a split pot is possible and explain in detail how the payout-weighted probabilities are calculated. If no split pot is possible ex-ante ($\mu = 0$), which is true for the majority of hands, both lotteries are equivalent ($L = \tilde{L}$). Moreover, the probability of a split pot is in every showdown the same for the player that faces the right-skewed and the player that faces the left-skewed risk, so it should not systematically confound our elicited skewness effects.

Notably, there are situations where the likelihood of a split pot is equal to one ($\mu = 1$), where the hand will result in a split pot irrespective of the remaining cards drawn from the deck. In these situations, the weighted probability we observe is equal to 0.5, but no insurance option was offered as the hand outcome is deterministic. In our data set, we cannot perfectly distinguish between such situations and situations with a weighted probability of 0.5 that resulted in a split pot, but where the result was not deterministic, and thus insurance was offered. To address this issue, we exclude all observations with an observed probability of 0.5. Most of these observations refer to a situation where the underlying lottery is degenerate, and no insurance option was offered.[23] Moreover, a winning probability of 0.5 implies that the underlying risk is not skewed, but symmetric. The main results are unchanged if we include all observations or

[22] As players' best five-card hand includes precisely two of their private cards and three of the community cards, splits are comparably rare in Omaha Poker cash games compared to other Poker games. This is one reason why it is advantageous to focus on this particular poker variant in our study.

[23] This is implied by the following observation: In total, there are 111,521 observations (out of 4,562,352) with an observed probability of 0.5, 95,415 of them result in a split pot. In these 95,415 cases, the insurance option was chosen in 830 cases (0.9%). In the 16,106 cases that did not result in a split pot, the insurance was chosen in 3,528 cases (21.9%).

only exclude observations with an observed probability of 0.5 that ultimately resulted in a split pot.

One further limitation is that, unlike in experimental studies, individuals do not choose between a binary risk and its expected value (the case we discussed in the preceding theoretical section) but have to pay a margin of 1% of the expected value in case the safe option is selected. The margin amount is small and affects all decisions equally, so the fee should not confound our results.

4 Results

4.1 Descriptives

We first provide a descriptive analysis of insurance choice frequencies for varying levels of skewness, but constant variance.[24]

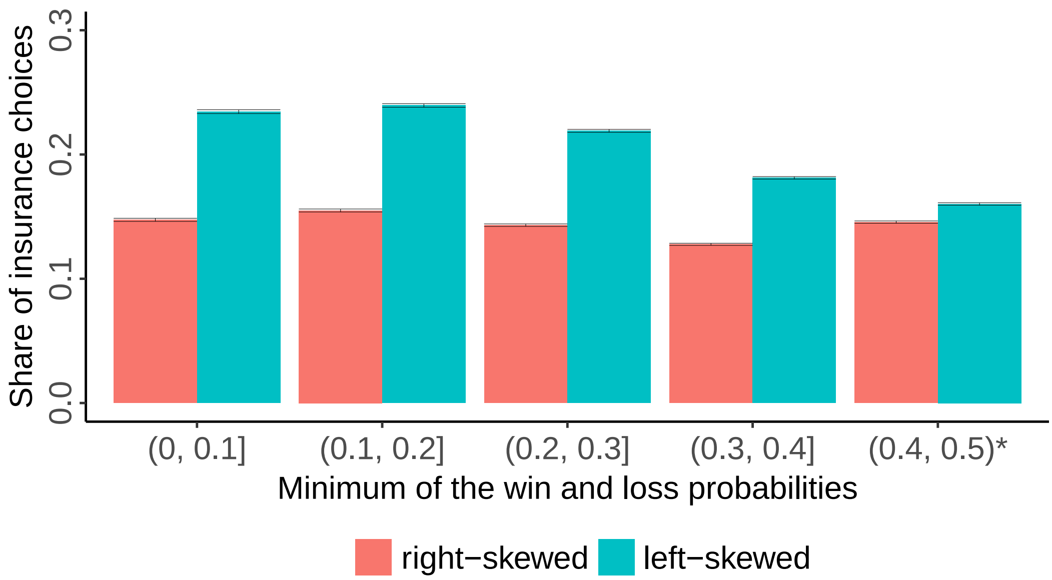
First, we group the observations conditional on the sign of skewness and find that individuals who face a negatively-skewed lottery choose the insurance option in 20.0% of cases. In contrast, individuals who face a positively-skewed risk do so in only 14.2% of the cases. This difference is highly statistically significant (p-value < 0.0001) and in line with a preference for positive skewness.

Second, Figure 2 illustrates the choice frequencies of the insurance option for different ranges of ex-ante winning probabilities with constant average variance. The horizontal axis depicts the winning probability range of right-skewed lotteries and the loss probability of the complementary left-skewed lotteries. For example, the first red bar at the left illustrates the insurance choice frequency for lotteries with a winning probability between 0 and 0.1, while the neighboring blue bar plots the frequency for lotteries with winning probabilities between 0.9 and 1, i.e., loss probabilities between 0 and 0.1. In all subgroups, individuals who face a negatively (left-)skewed lottery choose the insurance option significantly more often than their opponents who face a positively (right-)skewed lottery with the same variance. The differences are sizable and highly statistically significant in all groups and range from 1.5% to 8.7% (see Table 11, Appendix A.4). The differences are smallest for ranges closer to $\pi = 0.5$, i.e., lotteries that,

[24]Note that as we focus on situations where two players are in a showdown, i.e., winning the pot is a complementary event for both opponents, each right-skewed lottery ($\pi < 0.5$) has exactly one left-skewed lottery ($1 - \pi > 0.5$) as a complement. The variance of the opponents' binary lotteries are identical, $V_1 = V_2 = \pi(1 - \pi)x^2$. Accordingly, the opponent of a player facing a binary risk with skewness S faces a lottery with a skewness of $-S$.

on average, have smaller absolute skewness. Insurance shares tend to be smaller for lotteries with a larger variance, suggesting a positive preference for the variance for the average player. Interestingly, insurance shares seem rather constant concerning the variance for right-skewed lotteries and tend to decrease for a higher variance for left-skewed lotteries.

Figure 2: Share of insurance choices for different winning probability ranges



Note: Figure 2 depicts the share of insurance choices depending on ex-ante winning probabilities. The probability space is divided into 10 equidistant segments. Right-skewed and complementary left-skewed risks with the same variance are grouped together (e.g., the first red bar at the left refers to the interval of right-skewed risks with winning probabilities in the range $(0, 0.1]$ while the neighboring blue bar refers to the interval of left-skewed risks with loss probabilities in $(0, 0.1]$). For more details on observations and differences between groups see Table 11, Appendix A.4.

While these results support a preference for skewness, they do have certain drawbacks, as they, for instance, do not account for different expected values of the gambles and other factors that may confound our results. We address these issues in our regression analyses in the next sections.

4.2 Regression analyses

Empirical Strategy

We are interested in the the effects of each moment of the underlying probability distribution on individual insurance choice, keeping the other moments constant. We are agnostic about the underlying risk preferences model and follow Mitton and Vorkink (2007) in assuming that the different risk moments have a linear effect on utility. We do not have a clear prior regarding the influence of the expected value, given that both the safe option and the lottery exhibit roughly the same expected value. In contrast, we expect a positive (negative) sign for variance if individuals in our sample are, on average risk-averse (risk-seeking). Skewness preferences imply a negative skewness coefficient, meaning that individuals choose the risky option more often for higher skewness. In our main specifications, we estimate the following reduced-form equation:

$$y_{i,j(t,z)} = \beta_0 + \beta_E E_j + \beta_V V_j + \beta_S S_j + \gamma \mathbf{Z}_i + \eta \mathbf{W}_j + \lambda_t + \psi_z + \epsilon_{i,j}, \tag{5}$$

where the dependent variable $y_{i,j(t,z)}$ is a binary indicator of whether player i chooses the insurance option in decision j, that refers to a specific month t and game with stake z. Variables E_j, V_j and S_j denote the expected value, variance, and skewness of the binary risk in decision j. Variable λ_t gives the month fixed effects that control for month-specific factors constant across players that may affect risk-taking behavior, such as seasonality, adaptions over time, or COVID-19 effects, and ψ_z captures stake fixed effects that account for the fact that a game with a higher stake directly implies higher average expected values and variance. If such fixed effect controls are not included, our coefficients do not only capture the effect of E_j and V_j on insurance choice but also unobserved heterogeneity between games with different stakes. Finally, $\epsilon_{i,j}$ denotes the error term.

To account for confounding factors related to the features of a particular hand, we control for hand-specific characteristics $\mathbf{W}_j$, including the amount of money the player started the hand with (the stack), whether the player risked the entire stack during a particular showdown, the weekday, and the position of the respective player at the table.[25] The vector $\mathbf{Z}_i$ includes a set of player-specific characteristics to control for the playing style and the experience level of different players. Player-specific characteristics are based on all poker hands in our data set (including those without a showdown) and cover the following variables: number of hands played, num-

[25] Two remarks on the hand-specific characteristics: i) in a two-person showdown, there is always one player that is all-in, i.e., she risks all the money that she started the hand with (stack), because otherwise, betting between the two players would still be possible, which rules out a showdown; ii) the position at the table indicates when a player has to act during the hand, which may have important implications for the playing style and whether a player decides to play (i.e., voluntarily putting money in the pot) a particular starting hand or not. For example, players that already put money into the pot by posting the mandatory blinds tend to play a wider selection of starting hands, as the posted blinds count towards the necessary amount they have to call to see the first three community cards. Similarly, players who act last in each betting round (button) will have more information on opponents' actions when they act in future betting rounds, which usually increases the range of played starting hands as well.

ber of showdowns, profit or loss per 100 hands played, and the average winning probability at showdown over all hands (summary statistics can be found in Appendix A.4).

While including player-specific controls addresses some endogeneity concerns, our estimated coefficients may still be biased if unobserved factors are correlated with the type of lotteries individuals are facing. To address this issue, we exploit the panel structure of our data and include player-specific fixed effects α_i that control for all time-invariant heterogeneity across individuals. We extend the previous specification and also estimate the following fixed effect regression:

$$y_{i,j(t,z)} = \beta_0 + \beta_E E_j + \beta_V V_j + \beta_S S_j + \eta \mathbf{W}_j + \lambda_t + \psi_z + \alpha_i + \epsilon_{ij} \quad (6)$$

Regression results

Table 3 shows the estimated marginal effects from a linear probability model estimating equation (3) and (4). To compare the magnitude of the different coefficients more easily, we standardize the different moments in our main specifications.[26] The signs and the p-values of the coefficients are largely unchanged if we use non-standardized variables (see Table 12, Appendix A.5). Similarly, our main results are unchanged if we estimate a Probit or a Logit model instead (see Table 13, Appendix A.5). Using the standardized variables has the additional advantage that the coefficient of the constant can be approximately interpreted as the average insurance choice shares in the particular (sub)sample.

If we only include the three moments as independent variables (see Column 1, Table 3), we find that increasing skewness by one standard deviation leads to a 2.3 percentage point decrease in the likelihood that the insurance option is chosen, keeping the expected value and the variance constant. This is equivalent to a decrease of 13.4% compared to the average likelihood that the insurance option is chosen (i.e., the mean dependent variable, which is equivalent to the constant in our regression using z-scores). The coefficient is highly statistically significant (p-value<0.0001).

The coefficients of the other moments are significantly smaller. An increase of one standard deviation in variance, all else equal, decreases the probability that the insurance is chosen only by 0.04 percentage points, which is statistically insignificant at the 5%-level. Given that we are

[26]We follow the literature and standardize the variables by computing the z-score, that is, we subtract the respective mean and scale the variable by the inverse of its standard deviation. This allows us to make a unit-independent comparison of the coefficient magnitudes.

Table 3: Regression results for full sample

	Dependent variable:			
	Insurance choice dummy			
	(1)	(2)	(3)	(4)
Expected value	0.002***	0.003***	0.003***	0.007***
	(8.066)	(12.939)	(13.542)	(7.166)
Variance	−0.0004*	−0.001***	−0.001***	−0.003***
	(−1.689)	(−3.268)	(−2.876)	(−4.757)
Skewness	−0.023***	−0.023***	−0.024***	−0.023***
	(−124.166)	(−124.124)	(−125.872)	(−18.577)
Constant	0.171***	0.171***	0.171***	0.171***
	(965.613)	(974.794)	(966.475)	(3,676.599)
Player-specific controls	No	Yes	Yes	No
Hand-specific controls	No	No	Yes	Yes
Player fixed effects	No	No	No	Yes
Observations	4,450,585	4,449,739	4,449,739	4,450,585

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5). The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional control variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Column 4 provides results from the fixed effects regression (Equation 4) that includes both individual fixed effects of players and hand-specific controls. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns 2 and 3 differ because the average winning probability, one of the player-specific control variables, is not available in 846 choice situations. *: p<0.1; **: p<0.05; ***: p<0.01.

dealing with a sample of more than 4.4 million observations the insignificance of the second moment is particularly noteworthy. In comparison, the t-statistic of the skewness parameter is equal to −124.17. A higher expected value increases the probability that the insurance is chosen by 0.2 percentage points. Note that this does not mean that Poker players dislike positive returns because the insurance value also increases with the risk's expected value. The positive coefficient indicates that players are more likely to choose the insurance option when facing lotteries with larger expected values. While this comparison is only indicative, it is interesting to note that the absolute size of the skewness coefficient is more than 57 times larger than the variance coefficient.

In Column 2, we include player-specific characteristics as control variables. The estimated marginal effects of the expected value and skewness do not change considerably and remain highly statistically significant. The same holds true if we further include hand-specific control variables, see Column 3. The variance coefficient increases to 0.1 percentage points and becomes statistically significant at the 1%-level for these two specifications. However, the t-statistics and magnitude of the variance coefficient remains substantially smaller than for the other two moments, in particular compared to skewness.

The results of the player fixed effects regression model are presented in Column 4 of Table 3. The absolute coefficient of the first two moments increase to 0.7 percentage points for the expected value and -0.3 percentage points for the variance. The estimated effect of increasing skewness on the likelihood to choose the insurance option does not change in magnitude compared to the base specification. While the effect of variance is now significant at the 1%-level, it is still smaller in absolute terms than the coefficients on skewness, both in magnitude and significance.

In sum, our regression analyses reveal strong and statistically significant effects of skewness on individuals' decisions to take up the insurance option or not. Despite our very extensive data set, we fail to find a statistically detectable effect of variance—what is typically regarded as a risk's main property—in our base specification and only borderline statistically significant negative effects in the other specifications. The estimated effect of skewness is negative and strongly statistically and economically significant. The absolute magnitude and the t-statistics of the standardized skewness coefficients is considerably larger than for the other two moments, suggesting an preeminent role of skewness preferences for individual risk-taking.

4.3 Robustness of skewness preferences

In this subsection, we provide robustness checks to address potential endogeneity issues and limitations of our setup discussed above. First, as mentioned above, we estimate Logit and Probit models to account for the inherent non-linear relationship between our binary outcome variable and our independent variables. The results remain qualitatively robust. The estimated average marginal effects of skewness, as well as their significance, remain nearly unchanged in both specifications (see Table 13, Appendix A.5). For instance, in the most basic Logit and Probit specification, we estimate a skewness effect of -2.2 percentage points, compared to -2.3

percentage points in the linear probability model. While the coefficient on the expected value increases in both specifications, the variance coefficient is statistically insignificant for all Logit specifications (at the 1%-level) and only borderline significant and small in absolute terms for the Probit specifications (around -0.2 percentage points).

Second, the magnitude of the first two moments of the lotteries depend on the the net pot size. Our observations differ considerably in net pot sizes leading to substantial tails in the distribution of lotteries' expected value and variance. To make sure that this dispersion in pot sizes does not drive our results, we conduct our analyses using a normalized measure of lotteries' volatility, which is independent of the pot size: the "coefficient of variation" (CV) of the lotteries, which can be understood as the inverse of the "Sharpe ratio" of the lotteries.[27] The CV is defined as the ratio of the standard deviation to the mean: $\sqrt{Var(L)}/E(L) = \sqrt{\pi(1-\pi)x^2}/\pi x = \sqrt{(1-p)/p}$. This measure is dimensionless and commonly used in finance and economics (similar to the Sharpe ratio, Sharpe (1994)) and in psychology (e.g. Weber *et al.*, 2004). We estimate the same regression equation as above, interchanging the expected value and the variance of the lottery with its CV. Table 14 in the Appendix shows the estimated marginal effects. The skewness coefficient slightly increases to -2.6 to -2.4 percentage points, depending on the specification, and remains highly statistically significant. The coefficient on the CV is slightly positive in all specifications and statistically significant at the 1%-level (except for the specification with player fixed effects). As an additional corroboration that our results are not driven by outliers, we run our main specification with samples trimmed at the 1%- and the 99%-percentiles of the lotteries' net pot. The results are shown in Table 15. The skewness coefficient remains unchanged in main specification and slightly increases to -2.0 percentage points in the player fixed effects regression. The variance coefficient slightly increases in absolute terms to -0.5 percentage points and is now also highly significant. The coefficients are comparable if we trim the sample with respect to the expected at the 5%- and 95%-percentiles or if we winsorize the samples instead of trimming (results not shown).

Third, we run the same regressions for a sample that only includes players that face both types of lotteries—left- and right-skewed—at least once, which does not change the results (see Table 16, Appendix A.5). These results solidify our finding from above that the effects are not

[27]Note that the Sharpe ratio in Finance is usually defined in terms of the difference between a risky investment's return and the risk-free return.

driven by a fundamental difference between individuals facing left-skewed and individuals facing right-skewed risks.

Fourth, we estimate Equations 5 and 6, excluding all observations of players who never or always choose the insurance option. This rules out that our effect is driven by the fact that players who always choose the insurance option face fundamentally different binary risks than players who never choose the insurance option. Limiting our sample to those players increases the estimated effect size of all moments (in absolute terms). In the base specification, the estimated skewness effect increases from -2.3 to -3.0 percentage points, while the expected value and variance coefficients change to 1 and -0.4 percentage points respectively (Column 1 of Table 17, Appendix A.5). Similar changes can be observed for the other specifications (Columns 2-4). Again, the (absolute) effect size and t-statistics of the skewness variable largely exceed the estimates of the other moments in those specifications.

In response to the issue of split pots discussed in Section 3.3, we run the same regressions for a subsample that excludes all observations that resulted in a split pot. The results are illustrated in Table 18, Appendix A.5. The estimated coefficients and p-values of all three moments remain nearly unchanged compared to our main regressions (Table 3). Note, however, that this approach only excludes observations that ultimately resulted in a split pot and not all showdowns where a split pot is possible ex-ante. The robustness of our estimated effects reassures us that our results are not confounded by the possibility of split pots.

4.4 Heterogeneity in risk preferences

We estimate our main specifications for different subsamples to study heterogeneity in risk preferences and shed some light on the generalizability of our results. We split the full sample at the value of different player- and hand-specific characteristics that balances the observations in the two subsamples. [28] Skewness preferences can be observed and are robust in all subsamples, i.e., the coefficient of skewness is negative and strongly significant across all specifications that we consider.

[28] For a list and summary statistics of all characteristics, see Tables 6, 7, 8, 9 and 10, Appendix A.4.

Players' experience

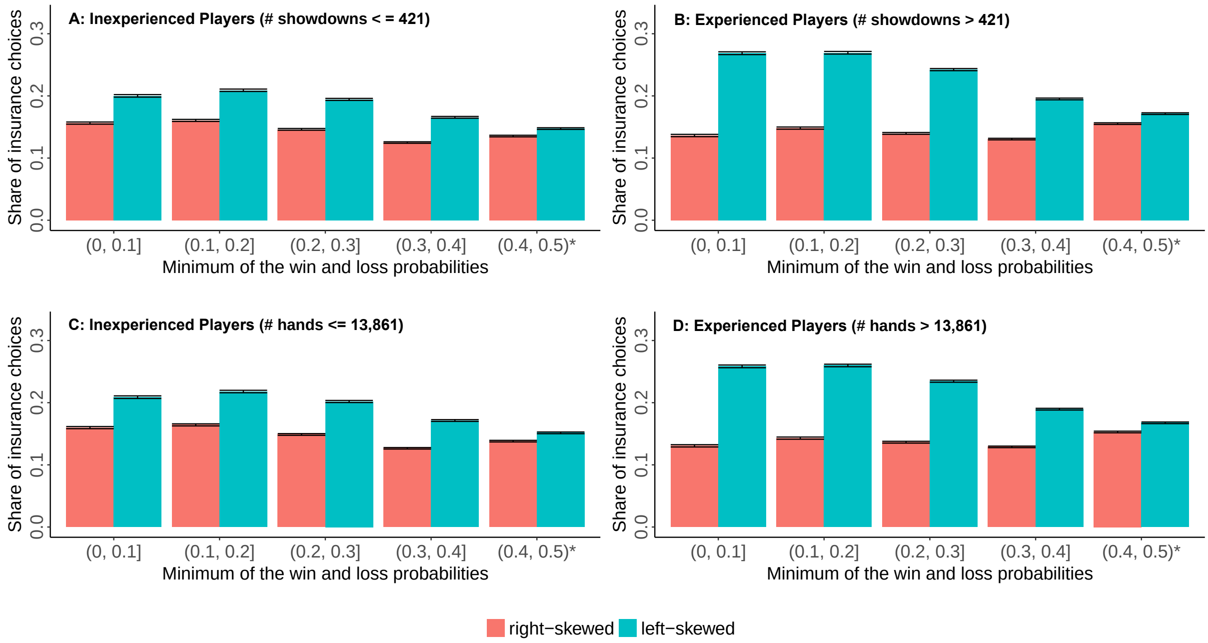
The experience level of the players is an important dimension. Previous studies already suggest that experienced poker players are more self-reflective, less affected by negative emotions, and make better decisions, by mathematical standards, than inexperienced players (Palomäki *et al.*, 2013, 2014). Therefore, for experienced players, the estimated effects are likely not driven by a misperception of the underlying lotteries or diluted by random and emotional choice but can rather be attributed to a deliberate choice of (or *preference* for) different moments of risk. Moreover, we have more repeated observations for more experienced players (for varying risks and payouts), which helps us to increase the power of our fixed effect regression model (see Table 19, Appendix A.6).

Experienced players face more than 421 showdowns, a definition that balances the observations in the subsamples of inexperienced and experienced players. Experienced players show a considerably stronger preference for skewness than inexperienced players. Experienced players choose the insurance option in 22.0% of the cases when facing a left-skewed lottery, compared to 14.2% when facing a right-skewed lottery. In contrast, the difference in insurance choice ratios between the left- and the right-skewed gamble is considerably smaller for the subsample of inexperienced players (17.9% vs. 14.2%). These differences along the experience dimension persist if we focus on varying levels of lotteries' variance (Panels A and B of Figure 3). Again, for both subgroups, differences in insurance choices between left- and right-skewed lotteries tend to be smaller the closer skewness is to zero (see Table 20, Appendix A.6).

In our regression analyses, we also find large heterogeneity in skewness preferences between experienced and inexperienced players. For experienced players (as measured via the number of showdowns), increasing skewness by one standard deviation decreases the likelihood of choosing the insurance option by 3.3 percentage points, which is equivalent to a decrease of 18.1% of the mean dependent variable in the respective subsample (Column 2, Table 4). For inexperienced players, the estimated effect of skewness on individuals' risk-taking is less than half in size (Column 1 Table 4). This also holds when we include individual fixed effects and hand-specific control variables (see Table 19, Appendix A.6).

The coefficients of the other two moments also differ considerably for experienced and inexperienced players. The effect of variance is quite small in all subsamples and has different signs for experienced and inexperienced players in the basic specification, similar to the expected

Figure 3: Share of insurance choices depending on players' experience



Note: The figure depicts the share of insurance choices depending on ex-ante winning probabilities, equivalent to Figure 2, for different subsamples of relatively inexperienced players (Panels A and C) and subsamples of relatively experienced players (Panels B and D). The sample is split at the median of the respective experience measure. In Panels A and B, players' experience is measured by the number of observed showdowns in the sample. In contrast, Panels C and D use the total number of played hands by each player, including those without a showdown, as a measure of the player's experience. The probability space is divided into 10 equidistant segments. Right-skewed and complementary left-skewed risks with the same variance are grouped together. For more details on observations and differences between groups see Table 20, Appendix A.6.

Table 4: Regression results for different levels of player experience

	Dependent variable:			
	Insurance choice dummy			
	# showdowns		# hands	
	≤ 421	> 421	≤ 13,861	> 13,861
	(1)	(2)	(3)	(4)
Expected Value	0.006***	−0.002***	0.007***	−0.002***
	(13.118)	(−8.310)	(12.494)	(−6.977)
Variance	−0.002***	0.001***	−0.003***	0.001***
	(−4.241)	(3.645)	(−4.117)	(3.837)
Skewness	−0.013***	−0.033***	−0.014***	−0.032***
	(−52.569)	(−115.220)	(−56.669)	(−111.701)
Constant	0.160***	0.182***	0.164***	0.178***
	(652.481)	(715.876)	(662.025)	(706.444)
Observations	2,228,808	2,221,777	2,225,785	2,224,800
Unique players	81,278	1,941	80,936	2,283

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5) for different subsamples. The full sample is split at the median of two different measures of player experience (as in Figure 3): i) the number of observed showdown situations per player (Columns 1 and 2); and ii) the total number of played hands by each player, including those without a showdown (Columns 3 and 4). At this place, we only present results from the specification without additional control variables or individual fixed effects of different players (equivalent to Column 1 of Table 3). The dependent variable is a binary indicator that equals one if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally, we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. *: $p<0.1$; **: $p<0.05$; ***: $p<0.01$.

value coefficient. These heterogeneity results further reaffirm our previous findings that skewness, compared to the variance, of the lotteries is a more stable and important driver of choice under risk in our sample.

In Panels C and D of Figure 3 and Columns 3 and 4 of Table 4 we split the sample according to an alternative measure of experience: the total number of hands played by an individual player (including those without a showdown). The cutoff value is equal to 13,861 played hands. The differences between the two subgroups are similar to before and remain essentially unchanged if we include fixed effects (see Table 19, Appendix A.6). The estimated coefficients of the other risk moments are also very similar for both experience measures.

Players' success

Skill partly drives poker players' success, which separates poker from other pure games of chance, such as roulette (Potter van Loon *et al.*, 2015; Duersch *et al.*, 2020). So, it may be relevant not to only distinguish between more and less experienced players but also between losing and winning players. While skill does not affect the outcome of the binary lotteries that we study in this paper, it affects the outcome of the preceding strategic interactions taking place under imperfect information. On average, skilled players more accurately judge the relative strength of their poker hands, predict the opponents' hands better and adjust their betting behavior accordingly, resulting in higher expected net returns. Therefore, there are likely systematic differences between losing and winning players concerning their playing motives, risk attitudes, and preferences.

We split the sample of players into two subgroups according to the recorded profits (also considering non-showdown hands). According to our definition, winning (losing) players make a positive (negative) net profit over our observation period. To account for the fact that profits can be the result of sheer luck if a player only plays a few hands, we also examine the heterogeneity in the subsample of experienced players only, who play at least 13,861 hands over our observation period. We distinguish between sophisticated players and recreational players. Sophisticated players are all experienced winning players, while experienced losing players are defined as recreational players. Like sophisticated individual investors, sophisticated poker players are likely to reflect and study their optimal (playing) strategy more frequently. Sophisticated players are less likely to play for recreational reasons only and are thus potentially more comparable to the financial experts. In contrast, recreational players and their underlying risk attitudes may be more comparable to speculative retail investors or people who gamble in a casino. Note that most winning players tend to be experienced. Over our observation period, the median experienced player makes an average loss of \$1.76 per hundred hands, while the median inexperienced player makes an average loss of \$20.83.

Table 5: Regression results in subsamples, depending on players' success

	Dependent variable:							
	Insurance choice dummy							
	All players				Experienced players only			
	without fixed effects		with fixed effects		without fixed effects		with fixed effects	
	Profit per hundred hands		Profit per hundred hands		Profit per hundred hands		Profit per hundred hands	
	≤ 0	> 0	≤ 0	> 0	≤ 0	> 0	≤ 0	> 0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expected Value	0.006***	0.002***	0.011***	0.003***	0.002***	0.002***	0.007***	0.002
	(14.164)	(8.058)	(5.599)	(3.343)	(5.591)	(5.586)	(2.690)	(1.636)
Variance	−0.002***	−0.001***	−0.004***	−0.002***	−0.0001	−0.001***	−0.002*	−0.001*
	(−3.605)	(−6.290)	(−3.512)	(−3.088)	(−0.560)	(−5.238)	(1.887)	(−1.690)
Skewness	−0.029***	−0.011***	−0.027***	−0.011***	−0.049***	−0.010***	−0.048***	−0.011***
	(−123.741)	(−40.649)	(−17.422)	(−7.172)	(−111.075)	(−34.107)	(−13.868)	(−5.264)
Constant	0.193***	0.114***	0.193***	0.114***	0.231***	0.105***	0.231***	0.105***
	(879.277)	(410.847)	(2,633.329)	(1,664.233)	(1,462.782)	(340.693)	(2,558.703)	(1,264.016)
Player-specific controls	No	No	No	No	No	No	No	No
Hand-specific controls	No	No	Yes	Yes	No	No	Yes	Yes
Player fixed effects	No	No	Yes	Yes	No	No	Yes	Yes
Observations	3,192,498	1,258,087	3,192,498	1,258,087	1,289,004	935,796	1,289,004	935,796
Unique players	65,886	17,333	65,886	17,333	1,472	811	1,472	811

Note: The table reports regression coefficients for basic OLS specification (Equation 5) in the columns (1)-(2), (5)-(6) and for the setup including player fixed effects (Equation 6) in columns (3)-(4), (7)-(8). The baseline samples are split into losing (columns with an uneven number) and winning (columns with an even number) players. Columns (1)-(4) employ the full sample and columns (5)-(8) the sample of experienced players as the baseline sample. The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. *: p<0.1; **: p<0.05; ***: p<0.01.

The estimation results of the subsample analyses are shown in Table 5. We find strong evidence of skewness preferences in both subsamples. However, skewness prefences are considerably more pronounced for losing players than for winning players, particularly if we consider the subsample of recreational players. For recreational players, a one standard deviation increase in skewness decreases the likelihood that the insurance is chosen by around five percentage points in the base case (Column 5) as well as in the fixed effects specification (Column 7). The magnitude and the size of the coefficents of the other two moments is comparable to the full sample. For winning and sophisticated players (Column 2, 4, 6 and 8), a one standard deviation increase in skewness decreases the likelihood that the insurance is chosen by only $1.0-1.1$ percentage points. While the estimated skewness effect is considerably smaller than for losing players, it is still highly significant in all specification in contrast to the other two moments. The absolute size of the expected value and variance coefficients are considerably smaller than for losing players, and become insignificant at the 5%-level for the fixed effect regression for sophisticated players (Column 8).

The subsamples do not only differ in terms of the estimated skewness coefficients, but also regarding the average insurance shares. Winning players choose the insurance in around 11.4% and losing players in around 19.3% of the cases. The difference in insurance shares increases if we consider recreational and sophisticated players (23.1% vs. 10.5%). The differences, both in terms of insurance shares and risk preferences, are intuitive in the sense that risk neutrality arguably helps to maximize earnings in the long run. Therefore, successful poker players' choices should be less sensitive to risk moments, and risk neutral players would never choose a costly insurance option. Yet, skewness still plays a significant role in the insurance decision of winning players.

5 Conclusion

The introduction of the insurance option in online poker allows us to cleanly test for skewness preferences in a large set of observational data among individuals that are rather experienced in choice under risk. We detect a strong and robust effect of skewness on risk-taking. Our results complement, for instance, recent survey findings (Holzmeister *et al.*, 2020) whereby skewness is the only moment that systematically affects financial professionals' perception of financial risk. We substantiate this finding in a real-world setting with a comprehensive data set of strongly incentivized stylized investment decisions. Our results suggest that idiosyncratic skewness or lottery-like features are important for asset prices, particularly in markets predominately populated by speculative investors, such as the markets for crypto assets. In addition, our findings have important real-world implications beyond asset pricing, as skewness preferences affect career choices and may explain recent phenomena, such as the boom of tech start-ups.

References

ALLAIS, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica: Journal of the Econometric Society*, **21**, 503–546.

ANDRIKOGIANNOPOULOU, A. and PAPAKONSTANTINOU, F. (2020). History-dependent risk preferences: Evidence from individual choices and implications for the disposition effect. *Review of Financial Studies*, **33** (8), 3674–3718.

BALI, T. G., CAKICI, N. and WHITELAW, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, **99** (2), 427–446.

BARBERIS, N. (2012). A model of casino gambling. *Management Science*, **58** (1), 35–51.

— and HUANG, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review*, **98** (5), 2066–2100.

— and XIONG, W. (2009). What drives the disposition effect? an analysis of a long-standing preference-based explanation. *Journal of Finance*, **64** (2), 751–784.

BARRAULT, S. and VARESCON, I. (2016). Online and live regular poker players: Do they differ in impulsive sensation seeking and gambling practice? *Journal of Behavioral Addictions*, **5** (1), 41–50.

BARSEGHYAN, L., MOLINARI, F., O'DONOGHUE, T. and TEITELBAUM, J. C. (2013). The nature of risk preferences: Evidence from insurance choices. *American Economic Review*, **103** (6), 2499–2529.

BELL, D. E. (1982). Regret in decision making under uncertainty. *Operations Research*, **30** (5), 961–981.

BERKHOUT, P., HARTOG, J. and WEBBINK, D. (2010). Compensation for earnings risk under worker heterogeneity. *Southern Economic Journal*, **76** (3), 762–790.

BORDALO, P., NICOLA GENNAIOLI and SHLEIFER, A. (2012). Salience theory of choice under risk. *Quarterly Journal of Economics*, **127** (3), 1243–1285.

—, — and SHLEIFER, A. (2013). Salience and asset prices. *American Economic Review, Papers & Proceedings*, **103** (3), 623–628.

Boyer, B., Mitton, T. and Vorkink, K. (2010). Expected idiosyncratic skewness. *Review of Financial Studies*, **23** (1), 169–202.

Boyer, B. H. and Vorkink, K. (2014). Stock options as lotteries. *Journal of Finance*, **69** (4), 1485–1527.

Brunnermeier, M. K. and Parker, J. A. (2005). Optimal expectations. *American Economic Review*, **95** (4), 1092–1118.

Chunhachinda, P., Dandapani, K., Hamid, S. and Prakash, A. J. (1997). Portfolio selection and skewness: Evidence from international stock markets. *Journal of Banking & Finance*, **21** (2), 143–167.

Cohen, A. and Einav, L. (2007). Estimating risk preferences from deductible choice. *American Economic Review*, **97** (3), 745–788.

Conrad, J., Dittmar, R. F. and Ghysels, E. (2013). Ex ante skewness and expected stock returns. *Journal of Finance*, **68** (1), 85–124.

Dertwinkel-Kalt, M., Ebert, S. and Köster, M. (2021). On correlated lotteries in economic applications. *Available at SSRN*.

— and Köster, M. (2020). Salience and skewness preferences. *Journal of the European Economic Association*, **18** (5), 2057–2107.

Duersch, P., Lambrecht, M. and Oechssler, J. (2020). Measuring skill and chance in games. *European Economic Review*, **127**, 103472.

Ebert, S. (2015). On skewed risks in economic models and experiments. *Journal of Economic Behavior & Organization*, **112**, 85–97.

— and Karehnke, P. (2020). Skewness preferences in choice under risk. *Available at SSRN*.

— and Wiesen, D. (2011). Testing for prudence and skewness seeking. *Management Science*, **57** (7), 1334–1349.

Eil, D. and Lien, J. W. (2014). Staying ahead and getting even: Risk attitudes of experienced poker players. *Games and Economic Behavior*, **87**, 50–69.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance*, **47** (2), 427–465.

Fiedler, I. C. and Rock, J.-P. (2009). Quantifying skill in games—theory and empirical evidence for poker. *Gaming Law Review and Economics*, **13** (1), 50–57.

Fischbacher, U., Hoffmann, G. and Schudy, S. (2017). The causal effect of stop-loss and take-gain orders on the disposition effect. *Review of Financial Studies*, **30** (6), 2110–2129.

Garrett, T. A. and Sobel, R. S. (1999). Gamblers favor skewness, not risk: Further evidence from united states' lottery games. *Economics Letters*, **63** (1), 85–90.

Gertner, R. (1993). Game shows and economic behavior: risk-taking on" card sharks". *Quarterly Journal of Economics*, **108** (2), 507–521.

Golec, J. and Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of Political Economy*, **106** (1), 205–225.

Green, T. C. and Hwang, B.-H. (2012). Initial public offerings as lotteries: Skewness preference and first-day returns. *Management Science*, **58** (2), 432–444.

Grove, W. A., Jetter, M. and Papps, K. L. (2021). Career lotto? labor supply in a superstar market. *Journal of Economic Behavior & Organization*, **183**, 362–376.

Gul, F. (1991). A theory of disappointment aversion. *Econometrica*, **59**, 667–686.

Hackethal, A., Hanspal, T., Lammer, D. M. and Rink, K. (2022). The characteristics and portfolio behavior of bitcoin investors: Evidence from indirect cryptocurrency investments. *Review of Finance*, **26**, 855–898.

Hartog, J. and Vijverberg, W. P. (2007). On compensation for risk aversion and skewness affection in wages. *Labour Economics*, **14** (6), 938–956.

Harvey, C. R. and Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance*, **55** (3), 1263–1295.

Holzmeister, F., Huber, J., Kirchler, M., Lindner, F., Weitzel, U. and Zeisberger, S. (2020). What drives risk perception? A global survey with financial professionals and laypeople. *Management Science*, **66** (9), 3977–4002.

Jondeau, E., Zhang, Q. and Zhu, X. (2019). Average skewness matters. *Journal of Financial Economics*, **134** (1), 29–47.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47** (2), 263–291.

Kraus, A. and Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *Journal of Finance*, **31** (4), 1085–1100.

Kumar, A. (2009). Who gambles in the stock market? *Journal of Finance*, **64** (4), 1889–1933.

Lin, T.-C. and Liu, X. (2018). Skewness, individual investor preference, and the cross-section of stock returns. *Review of Finance*, **22** (5), 1841–1876.

Loomes, G. and Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal*, **92** (368), 805–824.

Markowitz, H. (1952). The utility of wealth. *Journal of Political Economy*, **60** (2), 151–158.

Mitton, T. and Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *Review of Financial Studies*, **20** (4), 1255–1288.

Moskowitz, T. J. (2021). Asset pricing and sports betting. *Journal of Finance*, **76** (6), 3153–3209.

— and Vissing-Jørgensen, A. (2002). The returns to entrepreneurial investment: A private equity premium puzzle? *American Economic Review*, **92** (4), 745–778.

Palomäki, J., Laakasuo, M. and Salmela, M. (2013). "Don't worry, it's just Poker!"-Experience, self-rumination and self-reflection as determinants of decision-making in on-line Poker. *Journal of Gambling Studies*, **29** (3), 491–505.

—, — and Salmela, M. (2014). Losing more by losing it: Poker experience, sensitivity to losses and tilting severity. *Journal of Gambling Studies*, **30** (1), 187–200.

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

Potter van Loon, R. J., van den Assem, M. J. and van Dolder, D. (2015). Beyond chance? the persistence of performance in online poker. *PLOS one*, **10** (3), e0115479.

Prakash, A. J., Chang, C.-H. and Pactwa, T. E. (2003). Selecting a portfolio with skewness: Recent evidence from us, european, and latin american equity markets. *Journal of Banking & Finance*, **27** (7), 1375–1390.

Sharpe, W. F. (1994). The Sharpe ratio. *Journal of Portfolio Management*, **21**, 49–59.

Smith, G., Levere, M. and Kurtzman, R. (2009). Poker player behavior after big wins and big losses. *Management Science*, **55** (9), 1547–1555.

Snowberg, E. and Wolfers, J. (2010). Explaining the favorite–long shot bias: Is it risk-love or misperceptions? *Journal of Political Economy*, **118** (4), 723–746.

Sydnor, J. (2010). (Over)insuring modest risks. *American Economic Journal: Applied Economics*, **2** (4), 177–199.

Thaler, R. H. and Johnson, E. J. (1990). Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice. *Management Science*, **36** (6), 643–660.

Trautmann, S. T. and van de Kuilen, G. (2018). Higher order risk attitudes: A review of experimental evidence. *European Economic Review*, **103**, 108–124.

Vereshchagina, G. and Hopenhayn, H. A. (2009). Risk taking by entrepreneurs. *American Economic Review*, **99** (5), 1808–30.

Weber, E. U., Shafir, S. and Blais, A.-R. (2004). Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. *Psychological Review*, **111** (2), 430–445.

A Appendix

A.1 Poker hands ranking

The player with the highest ranked five-card hand, consisting of two private cards and three community cards, wins the pot in Omaha Poker cash games. The poker hand ranking is as follows (Source: Pokerstars: https://www.pokerstars.eu/poker/games/rules/hand-rankings/):

1. Straight Flush: Five cards in numerical order, all of identical suits. In the event of a tie, the highest rank at the top of the sequence wins. The best possible straight flush is known as a royal flush, which consists of the ace, king, queen, jack, and ten of a suit. A royal flush is an unbeatable hand.

2. Four of a Kind: Four cards of the same rank, and one side card or 'kicker.' In the event of a tie, the highest four of a kind wins. In community card games, where players have the same four of a kind, the highest fifth side card ('kicker') wins.

3. Full House: Three cards of the same rank, and two cards of a different, matching rank. In the event of a tie, the highest three matching cards wins the pot. In community card games, where players have the same three matching cards, the highest value of the two matching cards wins.

4. Flush: Five cards of the same suit. In the event of a tie, the player holding the highest ranked card wins. If necessary, the second-highest, third-highest, fourth-highest, and fifth-highest cards can be used to break the tie. If all five cards are the same rank, the pot is split. The suit itself is never used to break a tie in poker.

5. Straight: Five cards in sequence. In the event of a tie, the highest ranking card at the top of the sequence wins. Note: The Ace may be used at the top or bottom of the sequence, and is the only card which can act in this manner. A,K,Q,J,T is the highest (Ace high) straight; 5,4,3,2,A is the lowest (Five high) straight.

6. Three of a kind: Three cards of the same rank, and two unrelated side cards. In the event of a tie, the highest ranking three of a kind wins. In community card games, where players have the same three of a kind, the highest side card, and if necessary, the second-highest side card wins.

7. Two pair: Two cards of a matching rank, another two cards of a different matching rank, and one side card. In the event of a tie: Highest pair wins. If players have the same highest pair, highest second pair wins. If both players have two identical pairs, highest side card wins.

8. One pair: Two cards of a matching rank, and three unrelated side cards. In the event of a tie, the highest pair wins. If players have the same pair, the highest side card wins, and if necessary, the second-highest and third-highest side card can be used to break the tie.

9. High card: Any hand that does not qualify under a category listed above. In the event of a tie, the highest card wins, and if necessary, the second-highest, third-highest, fourth-highest, and smallest card can be used to break the tie.

A.2 Choice situation for a showdown before the flop

Figure 4 shows another example of a showdown situation, in particular of a constellation before any community cards have been dealt. The difference from the example in the main text is the different stage of the game when the showdown situation has occurred. The situation environment for the player is equivalent, i.e., payouts and probabilities are clearly displayed on the player's screen and the decision the player takes is the same. Note that the players have the insurance option only once, namely in the moment of showdown.

Figure 4: Example of an all-in cashout situation before any community cards have been dealt (German software)

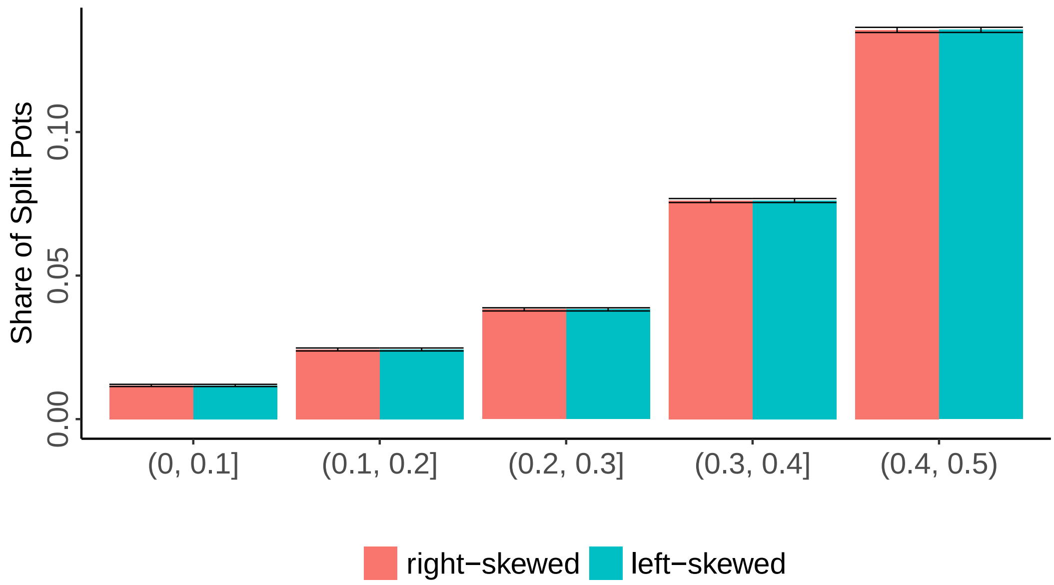
A.3 Showdowns with split pot possibility

Figure 5 shows an example of a showdown situation with one card to come. The difference from the example in the main text is that there is a split pot possibility. After the Turn, Player 1 holds the best five-card hand with a "Straight" (5-6-7-8-9). Player 4 best possible five-card hand is 8-8-8-K-9, three of a kind. Again, there are still 40 cards in the deck. Player 4 would win the entire pot if the board pairs, i.e., if a King, 9, 5 or 8 is drawn, giving her a winning Full House. As Player 1 holds one 8 and one 5, there are seven cards in the remaining deck that would give Player 4 the winning hand. Accordingly, the likelihood for Player 4 to win the entire pot is $\pi = \frac{7}{40} = 0.175$. However, as player 4 also holds a 6 (and 8/9) in his hand, a 7 on the River would give her the same straight (5-6-7-8-9) as Player 1, which would result in a split pot. There are three 7s still in the deck, implying a probability for a split pot of: $\mu = \frac{3}{40} = 0.075$. As Player 4 would only win half of the pot in this case, $\frac{1}{2}\mu$ is added to the winning probability to get the "payout-weighted probability," or expected winning share of pot, that is displayed on the screen: $\pi + \frac{1}{2}\mu = 0.175 + \frac{1}{2}0.075 = 0.2125$. Apart from that, the decision environment for the player is equivalent, i.e., payouts are clearly displayed on the player's screen and the players only have the insurance option once, namely in the moment of showdown.

Figure 6 illustrates shares of hands that result in a split pot depending on expected winning shares of the pot.

Figure 5: Example of an all-in cashout situation on the Turn with a split pot possibility

Figure 6: Share of split pots depending on the expected winning share of the pot



Note: Figure 6 presents the share of hands that resulted in a split pot depending on expected winning shares of the pot. The expected winning share is equivalent to the "payout-weighted" winning probability introduced in Section 3.3. The shares are divided into 10 equidistant segments and right-skewed and complementary left-skewed risks are grouped together (e.g., the first red bar on the left refers to the interval of right-skewed risks with expected winning shares in the range (0,0.1] and the neighboring blue bar refers to the interval of left-skewed risks with expected winning shares in [0.9,1)).

A.4 Additional summary statistics and descriptives

Table 6: Summary statistics of player-specific characteristics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Number of hands played	83,219	2,159.42	12,932.65	1.00	53.00	186.00	785.00	602,706.00
Number of experienced showdown situations	83,219	53.48	252.60	1.00	2.00	7.00	26.00	13,930.00
Average winning probability	82,470	0.44	0.15	0.00	0.37	0.45	0.52	1.00
Profit per hundred hands	83,219	−98.01	1,048.16	−145,158.40	−66.45	−19.73	−2.27	36,620.00

Note: The table reports summary statistics of all player-specific characteristics that we use in our empirical analysis. Characteristics are used, both, as control variables in our regressions analyses and to split the full sample for our subsample analyses (Section 4.3.). The variable *Number of experienced showdown situations* is calculated using the 4,485,585 observations of showdown situations in our full sample. The other characteristics are based on all hands in our initial data set, which also includes hands that did not result in a two-person showdown with a winning probability of $\neq 0.5$. The statistics are calculated with equal weights on all players. For the profit per hundred hands variable this implies that players who have played fewer hands are heavily overweighted in the calculation of the summary statistics. As these players usually make high losses in few hands and then stop playing, we get a large discrepancy between the average profits per hundred hand across all hands and players, which is equal to $-\$8.88$, and the profit per hundred hands when the mean is calculated with equal weights on single players (as illustrated in the table). The number of observations (N) for average winning probability differs compared to other characteristics as these values are not available in the data for 749 players.

Tables 7-9 report summary statistics of all hand-specific characteristics that we use in our empirical analysis.

Table 7: Summary statistics of stakes & stacks

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Stake	4,450,585	1.195	5.505	0.100	0.100	0.250	0.500	100.000
Stack	4,450,585	115.088	628.627	0.100	10.390	25.240	63.960	81,643.93

Note: The table reports summary statistics of the stake (mandatory bets) and the stacks (money of each player at the beginning of the hand) in our sample. Values are measured in US-$ terms.

Table 8: Summary statistics of risk-all-stack dummy

Statistic	N	Risk-all-stack=1	Risk-all-stack=0	Mean	St. Dev.
Risk-All-Stack Dummy	4,450,585	2,225,425	2,225,160	0.500	0.500

Note: The table reports the number of hand situations where the respective player risks her entire stack (Risk-all-stack=1). As we do not directly observe whether a player risks her entire stack in a showdown, we approximate this indicator variable based on the player's stack, final pot and expected winning shares. Due to rounding and presence of mandatory bets of other players who are not involved in the showdown, there might be some individuals that wrongly end up in the subsample of individuals that do not risk their entire stack. The error margin should be small and should not confound the results.

Table 9: Frequencies of table positions

	BB	BTN	CO	EP	MP	SB
Frequency	915,826	902,579	773,689	362,749	632,235	863,507

Note: The table reports absolute frequencies of the different table positions of the player in a showdown situation. These are namely: BB ("Big Blind"; person that has to post the big blind), BTN ("Button"; person that acts last in every betting round after the Flop), CO ("Cut-Off"; person that acts second last in every betting round after the Flop), SB ("Small Blind", person that has to post the big blind) as well as EP ("Early Position") and MP ("Middle Position”) that refers to positions between the "Big Blind" and the "Cut-Off".

Table 10: Frequencies of weekdays

	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
Frequency	610,961	610,300	617,142	616,895	641,434	682,997	670,856

Note: The table reports absolute frequencies of the weekdays when showdown situations have occurred.

Table 11: Differences in shares of insurance choice among right- and left-skewed risks for different ranges of winning probabilities

right-skewed interval (r)	left-skewed interval (l)	Obs.(r)	Obs.(l)	Δ in shares	t-statistic
$(0.0, 0.1]$	$[0.9, 1.0)$	$298,984$	$298,984$	0.087***	86.165
$(0.1, 0.2]$	$[0.8, 0.9)$	$344,012$	$344,013$	0.085***	88.607
$(0.2, 0.3]$	$[0.7, 0.8)$	$462,970$	$463,002$	0.076***	95.299
$(0.3, 0.4]$	$[0.6, 0.7)$	$563,305$	$563,297$	0.054***	78.788
$(0.4, 0.5)$	$(0.5, 0.6)$	$555,986$	$556,032$	0.015***	21.340

Note: The table reports the number of observations in each of the ten equidistant probability ranges illustrated in Figure 2. Column 5 reports the difference in the shares of insurance choice between complementary groups of left- and right-skewed risks. The imbalance in observations between the left- and right skewed intervals is due to rounding differences in winning probabilities. Corresponding t-statistics are displayed in Column 6.***: p<0.0001.

A.5 Robustness checks

Table 12: Regression results for full sample, non-standardized variables

	Dependent variable:			
	Insurance choice dummy			
	(1)	(2)	(3)	(4)
Expected Value	0.000005***	0.000008***	0.000009***	0.00002***
	(8.066)	(12.939)	(13.542)	(7.166)
Variance	−0.0000000002*	−0.0000000003***	−0.0000000003***	−0.000000001***
	(−1.689)	(−3.268)	(−2.876)	(−4.757)
Skewness	−0.010***	−0.010***	−0.011***	−0.010***
	(−124.166)	(−124.124)	(−125.872)	(−18.577)
Constant	0.186***	0.064***	0.044***	0.139***
	(224.891)	(34.965)	(22.722)	(36.403)
Player-specific controls	No	Yes	Yes	No
Hand-specific controls	No	No	Yes	Yes
Player fixed effects	No	No	No	Yes
Observations	4,450,585	4,449,739	4,449,739	4,450,585

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5). Compared to Table 3, the independent variable enters the regression as non-standardized absolute values. The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and over time (on the month level) to control for the unobserved heterogeneity in these dimensions. Corresponding t-statistics are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional controls variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Column 4 provides results from the fixed effects regression (Equation 4) that includes both individual fixed effects and hand-specific controls. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns (2) and (3) differ because the average winning probability, one of the player-specific control variables, is not available in 846 choice situations. *: p<0.1; **: p<0.05; ***: p<0.01.

Table 13: Regression results for full sample, Logit and Probit specifications

	Dependent variable:					
	Insurance choice dummy					
	Logit			Probit		
	(1)	(2)	(3)	(1)	(2)	(3)
Expected value	0.011^{***}	0.014^{***}	0.020^{***}	0.008^{***}	0.010^{***}	0.012^{***}
	(12.521)	(13.452)	(16.809)	(12.982)	(14.245)	(15.093)
Variance	-0.004^{**}	-0.006^{***}	-0.004^{*}	-0.002^{***}	-0.003^{***}	-0.002^{***}
	(−2.418)	(−2.811)	(−1.861)	(−3.410)	(−3.700)	(−3.788)
Skewness	-0.022^{***}	-0.022^{***}	-0.023^{***}	-0.023^{***}	-0.023^{***}	-0.023^{***}
	(−113.741)	(−112.777)	(−113.001)	(−117.179)	(−116.262)	(−116.751)
Player-specific controls	No	Yes	Yes	No	Yes	Yes
Hand-specific controls	No	No	Yes	No	No	Yes
Observations	4,450,585	4,449,739	4,449,739	4,450,585	4,449,739	4,449,739

Note: The table reports (average) marginal effects from estimating Logit & Probit specifications according to Equation 5. The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. The values of Wald test statistics (for testing the null hypothesis that coefficients are zero) are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional controls variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns 2 and 3 differ because the average winning probability, one of the player-specific control variables, is not available in 846 choice situations. *: $p<0.1$; **: $p<0.05$; ***: $p<0.01$.

Table 14: Regression results, employing coefficient of variation as a measure of dispersion instead of expected value and variance

	Dependent variable:			
	Insurance choice dummy			
	(1)	(2)	(3)	(4)
Coefficient of Variation	0.003***	0.003***	0.003***	0.0001
	(8.940)	(8.222)	(8.339)	(0.155)
Skewness	−0.026***	−0.026***	−0.026***	−0.024***
	(−68.267)	(−68.309)	(−69.277)	(−16.542)
Constant	0.171***	0.171***	0.171***	0.171***
	(965.620)	(974.794)	(968.129)	(2,502.023)
Player-specific controls	No	Yes	Yes	No
Hand-specific controls	No	No	Yes	Yes
Fixed effects	No	No	No	Yes
Observations	4,450,585	4,449,739	4,449,739	4,450,585

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5). The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the coefficient of variation and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional control variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Column 4 provides results from the fixed effects regression (Equation 4) that includes both individual fixed effects of players and hand-specific controls. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns 2 and 3 differ because the average winning probability, one of the player-specific control variables, is not available in 846 choice situations. *: p<0.1; **: p<0.05; ***: p<0.01.

Table 15: Regression results for a subsample with the net pot to be trimmed between 1%- and 99%-percentiles

	Dependent variable:			
	Insurance choice dummy			
	(1)	(2)	(3)	(4)
Expected value	0.009***	0.012***	0.012***	0.024***
	(32.338)	(43.577)	(43.526)	(14.446)
Variance	−0.005***	−0.006***	−0.006**	−0.009***
	(−26.117)	(−31.631)	(−30.242)	(−11.648)
Skewness	−0.022***	−0.022***	−0.022***	−0.020***
	(−114.681)	(−112.867)	(−114.551)	(−16.245)
Constant	0.173***	0.173***	0.174***	0.172***
	(961.479)	(970.495)	(936.394)	(1,942.803)
Player-specific controls	No	Yes	Yes	No
Hand-specific controls	No	No	Yes	Yes
Player Fixed effects	No	No	No	Yes
Observations	4,361,753	4,360,925	4,360,925	4,361,753

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5). The underlying sample is trimmed between 1%- and 99%-percentiles of the net pot. The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional control variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Column 4 provides results from the fixed effects regression (Equation 4) that includes both individual fixed effects of players and hand-specific controls. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns 2 and 3 differ because the average winning probability, one of the player-specific control variables, is not available in 828 choice situations. *: $p<0.1$; **: $p<0.05$; ***: $p<0.01$.

Table 16: Regression results for the subsample of players that face both left- and right-skewed showdown situations

	Dependent variable:			
	Insurance choice dummy			
	(1)	(2)	(3)	(4)
Expected Value	0.002^{***}	0.003^{***}	0.003^{***}	0.007^{***}
	(7.833)	(12.842)	(13.364)	(7.147)
Variance	−0.0003	-0.001^{***}	-0.001^{***}	-0.003^{***}
	(−1.588)	(−3.241)	(−2.888)	(−4.748)
Skewness	-0.023^{***}	-0.023^{***}	-0.024^{***}	-0.023^{***}
	(−124.566)	(−124.864)	(−126.641)	(−18.601)
Constant	0.171^{***}	0.171^{***}	0.171^{***}	0.171^{***}
	(962.681)	(972.125)	(963.279)	(3,665.002)
Player-specific controls	No	Yes	Yes	No
Hand-specific controls	No	No	Yes	Yes
Fixed effects	No	No	No	Yes
Observations	4,415,012	4,414,920	4,414,920	4,415,012

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5). Compared to Table 3, we exclude all observations of players that do not face at least one right- and one left-skewed lottery. The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional controls variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Column 4 provides results from the fixed effects regression (Equation 4) that includes both individual fixed effects and hand-specific controls. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns 2 and 3 differ because the average winning probability, one of the player-specific control variables, is not available in 92 choice situations. *: p<0.1; **: p<0.05; ***: p<0.01.

Table 17: Regression results for a subsample, excluding players who never or always choose the insurance option

	Dependent variable:			
	Insurance choice dummy			
	(1)	(2)	(3)	(4)
Expected Value	0.010***	0.011***	0.012***	0.017***
	(17.875)	(18.859)	(20.047)	(5.734)
Variance	−0.004***	−0.004***	−0.004***	−0.007***
	(−4.838)	(−5.147)	(−5.273)	(−4.072)
Skewness	−0.030***	−0.029***	−0.029***	−0.028***
	(−128.099)	(−122.344)	(−122.821)	(−18.616)
Constant	0.218***	0.218***	0.218***	0.218***
	(989.510)	(999.045)	(968.764)	(2,056.112)
Player-specific controls	No	Yes	Yes	No
Hand-specific controls	No	No	Yes	Yes
Fixed effects	No	No	No	Yes
Observations	3,476,736	3,476,692	3,476,692	3,476,736

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5). Compared to Table 3, we exclude all observations of players here who never or always choose the insurance option. The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional controls variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Column 4 provides results from the fixed effects regression (Equation 4) that includes both individual fixed effects and hand-specific controls. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns 2 and 3 differ because the average winning probability, one of the player-specific control variables, is not available in 44 choice situations. *: $p<0.1$; **: $p<0.05$; ***: $p<0.01$.

Table 18: Regression results for the subsample in which no split pots occur

	Dependent variable:			
	Insurance choice dummy			
	(1)	(2)	(3)	(4)
Expected Value	0.002***	0.003***	0.004***	0.008***
	(10.673)	(13.966)	(14.950)	(7.365)
Variance	−0.001**	−0.001***	−0.001***	−0.003***
	(−2.055)	(−3.056)	(−2.694)	(−4.127)
Skewness	−0.024***	−0.024***	−0.024***	−0.024***
	(−123.847)	(−123.122)	(−124.966)	(−18.625)
Constant	0.177***	0.177***	0.177***	0.177***
	(950.791)	(959.921)	(949.999)	(3,089.977)
Player-specific controls	No	Yes	Yes	No
Hand-specific controls	No	No	Yes	Yes
Fixed effects	No	No	No	Yes
Observations	4,154,930	4,154,123	4,154,123	4,154,930

Note: The table reports OLS regression coefficients of our main empirical specification (Equation 5). Compared to Table 3, we exclude all observations that result in a split pot ex-post. The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. Specifications (1)-(3) differ in the considered control variables: (1) does not contain any additional controls variables, (2) extends the basic specification by including player-specific characteristics as controls, (3) takes both player- and hand-specific characteristics into account. Column 4 provides results from the fixed effects regression (Equation 4) that includes both individual fixed effects and hand-specific controls. Underlying standard errors for the fixed effects regression are clustered at the individual level. Number of observations in Columns 2 and 3 differ because the average winning probability, one of the player-specific control variables, is not available in 807 choice situations. *: p<0.1; **: p<0.05; ***: p<0.01.

A.6 Additional subsample analyses

Table 19: Regression results in subsamples, with fixed effects

	Dependent variable:			
	Insurance choice dummy			
	# showdowns		# hands	
	≤ 421	> 421	≤ 13,861	> 13,861
	(1)	(2)	(3)	(4)
Expected Value	0.011***	0.003***	0.012***	0.003**
	(7.675)	(2.782)	(7.183)	(2.536)
Variance	−0.004***	−0.001**	−0.005***	−0.001**
	(−4.938)	(−2.217)	(−4.789)	(−1.959)
Skewness	−0.012***	−0.035***	−0.013***	−0.033***
	(−16.951)	(−14.646)	(−16.745)	(−14.183)
Constant	0.159***	0.182***	0.164***	0.178***
	(2,473.711)	(3,151.726)	(1,932.125)	(3,408.405)
Player-specific controls	No	No	No	No
Hand-specific controls	Yes	Yes	Yes	Yes
Player fixed effects	Yes	Yes	Yes	Yes
Observations	2,228,808	2,221,777	2,225,785	2,224,800
Unique players	81,278	1,941	80,936	2,283

Note: The table reports OLS regression coefficients of our fixed effects specification (Equation 6) for different subsamples. Compared to Table 4, we include individual fixed effects and hand-specific characteristics as control variables (equivalent to Column 4 of Table 3). The full sample is split at the median of two different measures of player experience: i) the number of observed showdown situations in our sample per player (Columns 1 and 2); and ii) the total number of played hands by each player, including those without a showdown (Columns 3 and 4). The dependent variable is a binary indicator that equals 1 if a player chooses the insurance option and zero otherwise. The main independent variables of interest are the expected value, variance, and skewness of the underlying lottery. Additionally we add fixed effects for different games (on the stake level) and different months to control for the unobserved heterogeneity across games and over time. The independent variables enter the regression as standardized z-scores. Corresponding t-statistics are provided in parentheses, using robust standard errors. *: p<0.1; **: p<0.05; ***: p<0.01.

Table 20: Differences in shares of insurance choice among right- and left-skewed risks

r	l	# showdowns		# hands	
		≤ 421	> 421	$\leq$ 13,861	$>$ 13,861
(1)	(2)	(3)	(4)	(5)	(6)
(0.0, 0.1]	[0.9, 1.0)	0.044***	0.132***	0.049***	0.128***
		(32.088)	(89.251)	(35.271)	(87.543)
(0.1, 0.2]	[0.8, 0.9)	0.049***	0.121***	0.054***	0.117***
		(37.326)	(86.976)	(40.776)	(84.803)
(0.2, 0.3]	[0.7, 0.8)	0.048***	0.103***	0.053***	0.098***
		(43.805)	(89.307)	(47.611)	(86.359)
(0.3, 0.4]	[0.6, 0.7)	0.041***	0.065***	0.045***	0.061***
		(42.663)	(66.379)	(46.302)	(63.359)
(0.4, 0.5)	(0.5, 0.6)	0.012***	0.016***	0.013**	0.015***
		(12.590)	(16.507)	(13.818)	(15.337)

Note: The table presents the difference in the shares of insurance choice between left- and right-skewed risks for different subsamples and ranges of ex-ante winning probabilities, as defined in Columns 1 and 2 and illustrated in Figure 3. The full sample is split at the median of two different measures of player experience: i) the number of observed showdown situations in our sample per player (Columns 3 and 4); and ii) the total number of played hands by each player, including those without a showdown (Columns 5 and 6). Corresponding t-statistics are provided in parentheses. *: p<0.1; **: p<0.05; ***: p<0.01.