Driven by Risk: Understanding Reference-Dependent Preferences using Simulated Auto Racing

November 26, 2024

- Submitted to JRI: July 3, 2024
- Manuscript ID: JRI-Jul-24-145
- express thanks to: ARIA, EGRIE, FSU Workshop, Johannes Jaspersen ...

Authors: James I. Hilliard* · Annette Hofmann† · David M. Weber‡

Abstract

Using data from over 56,000 simulated auto races worldwide, we analyze risk taking at the margins, consistent with reference-dependent preferences. We show that participants' risk-taking changes when a desired intermittent outcome is presented, sometimes at the expense of a more favorable expected end state. Specifically, we find that intermediate kinks in the reward function induce players to take reduce (increase) risk in the vicinity of opportunities to increase (lose) temporal wealth, providing important intuition regarding the incentives for risk-taking at the margin of wealth kinks (e.g. retirement age, family changes, etc.). Risk avoidance appears to increase as the kink approaches, but risk-taking increases with more individual investment.

*James I. Hilliard, Ph.D., is Associate Professor of Instruction at the Fox School of Business, Temple University, james.hilliard@temple.edu, (215) 204-8153.

†Annette Hofmann, Ph.D., (corresponding author) is Virgil M. Schwarm Associate Professor of Finance and Academic Director of the Lindner Center for Insurance and Risk Management at the Lindner College of Business, University of Cincinnati, hofmanae@ucmail.uc.edu, (347) 387-1976.

‡David M. Weber, Ph.D., is Senior Lecturer of Information Systems at Eller College of Management, University of Arizona, davidweber@arizona.edu, (520) 621-8649.

1 Introduction

Behavioral economics emerged as classical models of economic behavior were adjusted to anticipate observed behavior. For example, two paradoxes were introduced by Allais (1953b,a),[1] continued with a note by Samuelson (1963), and eventually stated as Prospect Theory by Kahneman and Tversky (1979). Today, Prospect Theory is widely accepted, recognizing that individual preferences may vary at certain reference points. For example, an individual may evaluate a risky outcome based on the wealth level achieved from previous choices and outcomes (Kőszegi and Rabin (2006)). The reference point divides between perceived gains and losses, resulting in behavior inconsistent with classical expected utility theory.

Given the Prospect Theory predictions about reference points, principal-agent theory predicts performance boosts in the final stages of long-term contracts and shirking in early stages of such contracts (O'Neill and Deacle (2019)). In particular, expected utility-maximizers can be assumed to exert effort until their marginal benefit equals marginal cost. The extent of moral hazard in this relationship depends on the individual's own costs of negotiating a new contract and the overall costs of monitoring individual performance (Fudenberg and Tirole, 1990). Thaler (1999) suggests that "mental accounting" assigns activities to various accounts and ignores the impact of outcomes of those activities on other activities. As a commonly observed phenomenon, mental accounting supports prospect theory in that agents will focus on one account to the detriment of another, and to the detriment of the agent's total portfolio.[2]

We find evidence consistent with the literature discussing final contract stages and their impact on individual decision-making. Increased individual effort in these final contract stages – "where it matters" – has been observed in economic studies in sectors such as sports and public procurement. Indeed, when a professional athlete approaches the final period of his contract and nears possible renewal, he balances the expected rewards of increased effort with the expected costs of resulting injuries – mentally optimizing the expected value of the next contract. Conversely, once the team extends the contract, the incentive to relax prevails, since early contract performance may matter less.

[1] It is shown that patterns of choice can be inconsistent with expected utility maximization models, mostly revolving around violations of the independence axiom within Expected Utility Theory.

[2] Notably, Rabin and Thaler (2001) suggest mental accounting as sufficient rationale for individuals carrying low liability limits while paying significant premiums on internal wiring insurance, inconsistent with expected utility maximization. Mental accounting would also justify varying effort over the life of a contract.

This is confirmed by Frick (2011) using the German Bundesliga for soccer player effort levels. When the outcome of individual effort is uncertain and risks are involved, individuals tend to increase risk-taking at the margin, aiming for potential better status, and then reducing risk-taking in other situations to protect their current status. Indeed, Frick (2011) finds that player performance lags, on average, in the early years of a player's contract and "significantly improves" in the last year of the contract, giving evidence of a meaningful kink in the reward structure.

This study answers such questions using a unique dataset containing over 56,000 simulated auto races. In auto racing simulations, participants join a virtual auto race, using home computers or gaming systems connected to a central server via the Internet. This study employs data from the simulation to evaluate whether and how individual risk taking changes when individuals get closer to a situation where performance matters more. It thereby contributes to the literature on risk taking by providing a rationale for adjustments in "final-contract-stage" behavioral patterns as related to individual risk taking at the margins.

Recent literature highlights that in field settings, it is not always obvious which reference points are important for subjects making decisions under risk. As discussed in Barberis (2013), identifying appropriate reference points is only possible in rare scenarios. One such scenario has been examined by Allen et al. (2017), who study a large data set of marathon finishing times and show that "round numbers" (including, for example, Boston Marathon qualifying times, or hour / half-hour increments) serve as reference points in such a running environment where runners may set round finishing times as their aspiration levels. Similarly, the scoring and advancement system of iRacing (our simulated racing platform) represents another ideal environment to study field evidence of such reference-dependent preferences. Thus, our results add to those of other studies where reference points may differ for individuals, and where the reference points are exogenously established by the platform rather than the individual. Since our large sample consists of (presumably) nearly exclusively male drivers, there is no gender effect to disturb risk taking choices. Another main difference of this study when compared to Allen et al. (2017) is that the marathon setting is not a zero-sum game; a runner's finishing time has no direct effect on another runner's finishing time (though, at elite levels where "place" matters, a runner's finishing position necessarily effects another runner – this is not binding on more casual racesr). Rather, a racer's success in iRacing negatively impacts all racers who finish behind a particular racer.

Does risk taking change when individuals get closer to a margin where "it matters more" how they perform? Following economic theory on principal-agent relationships, one would expect performance boosts in the final stages of long-term contracts and shirking in early stages of such contracts. In sports, for example, unpublished work by citetfeess2010incentive extends the work of Frick (2011), providing further empirical evidence that average performance of soccer players in the German Bundesliga decreases in contract length. Interestingly, long-term contracts impact the performance distribution asymmetrically, in the sense that they increase the probability of poor performances but do not reduce the probabilities of good performances. Iossa and Rey (2014) propose a theoretical explanation, finding that incentives are stronger and performance higher as a contract approaches renewal. In American professional basketball, for example, it has been shown that player performance improves significantly in the year before signing a multi-year contract, and declines after the contract has been signed (Stiroh, 2007). Similarly, O'Neill and Deacle (2019) provide evidence that Major League Baseball (MLB) players' effort levels vary according to their positions in the contract cycle, using 2007-2011 data. Specifically, by controlling for time-invariant player traits, fixed-effect regression models show that MLB players increases effort in the final years of their contracts. Players also tend to shirk in the first years of new long-term contracts that last four to six years.

In this vein, we extend the research to the (largely) non-professional e-sports world. Our simulated auto racing dataset provides evidence consistent with research about professional athletes that risk-taking behavior changes at the margins, where kinks in the reward function are present. While this setting does not have the pecuniary benefits associated with professional sports contracts, players are competing seriously for status (iRating) and access to more competitive heats (safety score, converted to license level, with potential for more significant iRating movement). Thus, we take advantage of an explicit kink in the safety score-license level relationship to explore risk-taking behavior at that kink.

This study is organized as follows: The next section reviews relevant literature. Section 3 presents the iRacing.com simulated auto racing platform and lays out basic characteristics of the gaming platform of interest. Section 4 describes the data and variables of interest used in the study. Section 5 derives our hypotheses. Section 6 presents the econometric model and empirical findings. The last section concludes.

2 Related Literature and Conceptual Framework

A study by Halek and Eisenhauer (2001) uses life insurance survey data to estimate the Pratt-Arrow coefficient of relative risk aversion for almost 2,400 households in the US. Attitudinal differences toward pure risk are examined across demographic subgroups. An interesting finding is that self-employed peoples' attitudes toward speculative risks do not differ from others, yet they are significantly more averse to downside risks than those employed by others. Similarly, at the margin, education increases individual risk aversion to pure risk but also increases the willingness to accept a speculative risk. This may relate to a desire to control one's environment: one must actively seek out speculative risks, whereas one merely reacts to pure risk when it is realized.

Brunnermeier and Nagel (2008) use the Panel Study of Income Dynamics to study how households' portfolio allocations change in response to wealth fluctuations. They find that persistent habits, consumption commitments, and subsistence levels can generate time-varying risk aversion, showing that the share of household liquid assets invested in risky assets is not affected by wealth changes. The authors identify inertia as one of the major drivers of household portfolio allocation. This finding is consistent with early research by Samuelson and Zeckhauser (1988), who show that agents are slow to make decisions about risk, tending to maintain prior decisions, even in the face of new information. Dohmen et al. (2011) find in a large study that personal characteristics have an economically significant impact on risk aversion in general. An additional experiment confirms the behavioral validity of this measure. Turning to other questions about risk attitudes in specific contexts, the authors find similar results on the determinants of risk attitudes, and also evaluate the potential lack of stability of risk attitudes across contexts. Peter (2021) generalizes several existing approaches in the literature, including Eeckhoudt and Gollier (2005) and Dionne and Li (2011), to obtain findings on how decision-makers trade off comparative risk aversion against comparative downside risk aversion. Whether both effects are aligned depends on a probability threshold unique to the benchmark agent's utility function, allowing for an entire class of decision-makers sharing the same comparative static prediction relative to this reference agent. Paravisini et al. (2017) estimate risk aversion from investors' financial decisions in a person-to-person lending platform and find that wealthier investors are more risk averse in the cross-section and that investors become more risk averse after a negative housing wealth shock.

Going beyond expected utility studies, Greer (1974) was the first of many questioning prevailing utility theory's ability to predict the decision-making process of individuals. Greer and Skekel (1975), following Hoskins (1975), argued that actual decision-making might result from a utility function of nonclassical shape – a function with one or more "kinks" – in contrast to the classical utility function attributed to von Neumann and Morgenstern.[3] These studies suggest that individuals may tend to be more risk-averse (i.e., the utility function might be more concave) up to a certain reference point or utility level, implying that risk aversion is not stable over final wealth levels. Indeed, empirical studies of kinked utility functions show that such utility functions exhibit first-order risk aversion at the kink and can explain the common empirical findings of a general preference for full insurance, even when the premium is not actuarially fair (Segal and Spivak (1990)).[4] Riley and Chow (1992) find that risk aversion tends to decline with wealth, level of education, and age, but only until the age of 65, when risk aversion goes up. Note that the traditional retirement age of 65 represents a "cusp" in the sense that agents adjust to a different lifestyle and income source.

In addition to the "cusps" we are faced with in our lives, our individual experience level and skill development may also impact our risk-taking. Chains of prior successes or failures can lead to changes in a person's risk attitude over time, despite the relegation of the "hot hand" to the pool of logical fallacies. In a study by Brocas et al. (2019), the impact of prior gains or losses was observed in the behavior of almost half of subjects. The majority took more risk after a gain. Hot hand fallacies notwithstanding, this phenomenon is often explained by the "House Money Effect" in which an agent plays as though money won in the game is separate from their total wealth. Similarly, it has also been shown in several experiments that people are less likely to take risks after prior losses. For instance, Kőszegi and Rabin (2006) predict reference-dependent risk attitudes, where reference points are determined endogenously by the economic environment. Here, the authors tested incremental outcomes with little understanding of prior decisions - prior tests relied on "house money". Since wealth is rarely known at the beginning of an experiment, the individual's decisions are only evaluated with respect to money that they earned inside the experiment.

Hersch and McDougall (1997) examine risky decisions as displayed in *Illinois Instant Riches*, a televised set of three different games of chance, providing a natural experiment for assessing

[3] See Greer and Skekel (1975), p. 843.

[4] Kinked utility is also discussed by Sinn (1982) in a different context.

the risk taking behavior of individuals. They find that income is insignificant as a predictor of risk taking. Post et al. (2008) study house-money based decisions in the game show *Deal or No Deal*. In contrast to traditional expected utility theory, they find that choices can be explained in large part by previous outcomes experienced during the game. Finally, Gertner (1993) studies the bonus round of the game show *Card Sharks*. He finds that asset segregation may be an important driver of decision-making under risk:

"If this phenomenon is present in the risky decisions that individuals make, such as insurance and portfolio selection, the expected-utility model may not be appropriate. People may make these decisions ignoring their overall wealth level, but instead, base the decision only upon the direct stakes involved."[5]

2.1 Conceptual Framework

This study evaluates risk taking by mostly amateur simulated auto racing participants, in particular "when it matters". It will be shown that when there are "cusps" or kinks in individual utility functions, following the idea of incentive wealth areas where effort matters, simulation participants take more risk given opportunities to increase wealth or status at a "cusp". Three examples of a functional form of such a utility function are shown in Figure 1. It can be demonstrated in a simple model that reference dependent preferences imply bunching of performance at or around the reference points, "where it matters". Indeed, reference dependent preferences entail that individuals assess outcomes marginally above or below the (neutral) reference point in a manner divergent from conventional utility theory. The perceptual differentiation of outcomes proximate to the reference point can manifest in various qualitatively distinct forms. For an individual with twice continuously differentiable utility $u(\cdot)_r$, using reference point r, reference dependence can take three primary forms:[6]

(a) Discontinuity in the form of a "jump" at r: $lim_{\epsilon \to 0}\ u_r(r+\epsilon) \neq lim_{\epsilon \to 0}\ u_r(r-\epsilon)$

(b) Discontinuity in the first derivative at r: $lim_{\epsilon \to 0}\ u'_r(r+\epsilon) \neq lim_{\epsilon \to 0}\ u'_r(r-\epsilon)$

(c) Discontinuity in the second derivative at r: $lim_{\epsilon \to 0}\ u''_r(r+\epsilon) \neq lim_{\epsilon \to 0}\ u''_r(r-\epsilon)$

[5]See Gertner (1993), p. 520.
[6]See Allen et al. (2017).

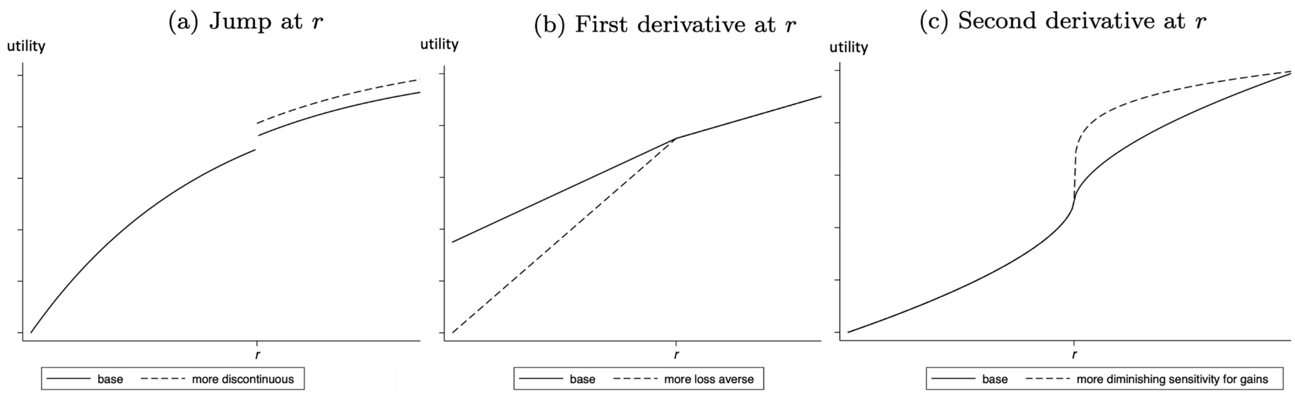

Figure 1: Utility functions with a cusp at r expressing different forms of reference-dependent preferences (Allen et al. (2017)).

The discontinuity in the first derivative at r is often referred-to as *loss aversion*. Wakker and Tversky (1993) define loss aversion as follows: An individual is loss averse if the utility or benefit function is everywhere steeper in losses than for the respective comparable gains. Then, function $u_r(.)$ expresses higher loss aversion than function $v_r(.)$ if the utility functions coincide for gains but $u_r(.)$ is steeper than $v_r(.)$ everywhere in losses (see (b)). It is shown by Allen et al. (2017) that "more loss averse" is related to bunching of performance near the reference point.[7] This is because higher loss aversion increases the marginal benefit of effort short of the reference point, boosting motivation to get there. Similarly, it can be shown that a discontinuity in the second derivative of the utility function, implying diminishing sensitivity to gains, can also lead to *bunching* near the reference point. However, any form of reference dependence will produce bunching and behavior will thus deviate from expected utility theoretical predictions.

[7] In economics, bunching often refers to the clustering of economic activities or behaviors around specific thresholds or notches created by policies or regulations. For example, if a tax policy imposes higher rates on income above a certain level, individuals might adjust their earnings to stay below that threshold, resulting in a bunching of incomes just below the cutoff point.

3 Description of the Racing Platform

3.1 eSports and Sim Racing

Starting in the early 1970s, eSports began as a niche hobby, with events such as a 1972 Stanford University Competition which involved the video game *Spacewar*. In this video game, participants competed for a "grand prize" of a *Rolling Stone* magazine subscription. Only recently, as of 2019, the eSports industry recorded an estimated \$1 billion revenue worldwide (Russ, 2019), with single event prize pools reaching \$25.5 million for the 2018 DOTA 2 International Competition (representing more than double the 2018 U.S. Open golf prize pool) (Ingraham, 2018). Single eSports events, like China's Battle of Balls Professional League, have attracted roughly 13,000 in-person viewers and 5 million online viewers (Chan, 2018). The City of Arlington, Texas is building a \$10 million, 100,000 square foot dedicated eSports arena at taxpayer expense (Albright, 2018). Lando Norris, a real-world Formula 1 driver, streams simulated races on Twitch and has 1.2 million followers.[8]

Sim racing is a genre of eSports that involves a user manipulating specialized hardware (steering wheels, pedals, shifters, and button boxes) to control a virtual car around a virtual track through simulator software running on a PC or gaming console. The driver sees the racetrack from the perspective of the vehicle's cockpit on their gaming monitors or through a virtual reality headset. A sample hardware setup for this platform is shown in Figure 2. Similar platforms exist for bicycle racing, running, and other endurance sports. The major advantage of using data from this software for research is that is makes the user experience very realistic and so conclusions drawn from our data should match real-world behavior.

Advancements in hardware, processing power, telemetry, and scanning methods have diminished the gap between sim racing and real world racing. Motoring media publisher TopGear reported that champion Finnish sim racer Greger Huttu, does not actually possess a driver's license and had never driven a real car[9]. Nonetheless, he came to the Road Atlanta facility and drove a 260bhp Star Mazda open-wheeler race car to determine how well he could translate his sim racing experience to the real world. The coaches were impressed with his speed. "The telemetry confirms it. His braking points are spot on. He's firm and precise on

[8] This Twitch channel can be accessed at `https://www.twitch.tv/landonorris` and metrics can be accessed at `https://www.twitchmetrics.net/c/174809651-landonorris`.

[9] This story is related at `https://www.topgear.com/car-news/gaming/geek-rebooted`.

Figure 2: Examples of Sim Racing hardware setups.

the throttle. And in the fastest corner, he's entering at 100mph compared to an experienced driver's 110 – a sign of absolute confidence and natural feel for grip. Remember, this is a guy who has never sat in a racing car in his life – he's only referencing thousands of virtual laps. Then, on lap four, he pops in a 1:24.8, just three seconds off a solid time around here." In another anecdote, a NASCAR team recruited William Byron based on his winning record in sim racing. Professional race car drivers, including Dale Earnhardt, Jr., Kyle Larson, and Joey Logano use sim racing to train and learn new tracks[10].

3.2 iRacing

iRacing (`http://www.iracing.com`), is a subscription-based sim racing platform developed by iRacing.com Motorsport Simulations in Bedford, MA and made available to the public in 2008. Co-founder John Henry is a principal owner of the Boston Red Sox and Liverpool Football Club. iRacing offers one of the market's most realistic and authentic simulated racing experiences to its members. They use LIDAR technology[11] to laser scan real-world racetracks which are accurate to one millimeter. For cars, iRacing developers use telemetry data and "scan, weigh, and measure each part of the actual race cars and assemble them digitally, giving users a mathematically correct vehicle to drive on the mathematically correct racing surface" (Brown, 2008).

[10] This story is related at `https://www.foxbusiness.com/features/nascar-taps-into-e-sports-to-recruit-young-fans-drivers`. Other user experiences are related at `http://iracing.com/testimonials/`.

[11] For a description of LIDAR, see `http://oceanservice.noaa.gov/facts/lidar.html`.

iRacing's platform is unique in that competition is solely on-line, in contrast to other racing simulation software (e.g., Assetto Corsa, Project Cars 2, Automobilista, rFactor 2). There is no single-player mode (with the exception of practice rounds, known as "hotlapping") and all competitors are human participants, with no artificial intelligence (AI) or random number generator (RNG) competition. All competition is against other subscribers to the service, which boasts over 200,000 members, 80 official series, 400 private leagues[12], and a 2021 annual prize pool of $330,000 for the eNASCAR Coca-Cola iRacing Series alone (see `https://www.enascar.com/coca-cola-iracing-series/`).

3.3 Pricing and Scoring Model

iRacing charges its participants a subscription fee, with premium content available for an additional charge. Its members, referred to as *iRacers*, subscribe by paying monthly or annual fees (ranging from $13 per month to $199 per two years) giving them access to iRacing's base content of 18 cars and 18 tracks. Additional cars (50+) and tracks (60+) are available to license at prices from $11.95 to $14.95, each with volume and package discounts available. iRacing has licensing partnerships with numerous real-world racing organizations and race car manufacturers and series (e.g., NASCAR, Renault, Mazda, Porsche, Blancpain, and McLaren[13]), which afford them the rights to model and sell vehicle licenses to iRacers.

To promote safe, organized, and competitive races, iRacing ranks each iRacer with two independent ratings: (1) iRating and (2) safety rating. An iRacer could have a very high safety rating with a low iRating or vice-versa. Races occur in 4 disciplines: Road, oval, dirt road (also known as rallycross), and dirt oval. iRacers have separate scores for iRating and safety rating in each category, but most iRacers focus on one racing category. The present study focuses on a series in the "road" discipline.

"iRating", following the Elo system used to rank chess players (Coulom, 2007), is a measure of an iRacer's ability to win races and finish near the front of the pack. Each race is a zero-sum game, where iRacers take a share of iRating from all drivers who finished behind them and lose iRating to all drivers who finished ahead of them. A driver who finishes ahead of another driver with a higher iRating will gain more iRating than beating a driver with

[12] Membership data current as of January 2022. A complete overview of the platform can be found at `http://iracing.com/overview/`.

[13] `https://www.iracing.com/partners/`.

a lower iRating. All iRacers start their career in each category at 1,350 iRating. Given the zero-sum game of iRating, that number is also the theoretical average iRating across all iRacers. In 2018, the highest iRating was held by Ty Majeski, a 22-year-old from Wisconsin whose iRacing success earned him a position as a development driver for Roush Fenway Racing (Lawrence, 2018). The highest iRating in mid-2022 was held by Max Benecke, a sim racer from Germany whose iRacing success earned him $46,000 in prize earnings in 2021[14] and he is a sim racing endurance event teammate of the 2021 real-world Formula 1 champion Max Verstappen.

"Safety rating" is a measure of a driver's ability to drive safely and avoid incidents, where incidents include leaving the track, losing control of the vehicle, or making contact with other cars. Each violation is considered an *incident,* with more severe violations penalized more harshly. For example, having more than 50% of one's car off the track surface results in a 1x incident, while a hard collision with another vehicle results in a 4x incident. Incidents do not assign blame, so in the case of a reckless driver who T-bones another car in the braking zone, both drivers would be assessed a 4x incident. Most races have an incident limit (for example, in most races lasting less than an hour, the incident limit is 17x). Once an iRacer surpasses that limit, the driver is immediately disqualified. For this study, a high incident count is assumed to be correlated with unobserved risk-taking behavior.

3.4 License Levels

iRacing uses the number of corners driven per incident received as the measure to determine a driver's safety rating. The iRacer sees their safety rating in terms of license classes. All iRacers begin in the rookie (R) license class and can progress through safe driving to D, C, B, A, and Pro licenses. Each level requires a progressively higher degree of safe racing and affords the iRacer opportunities to race in more challenging series. Similarly, license levels can be lost by amassing high incidents per corner. Once an iRacer crosses the threshold of a license level or sublevel (either an increase or decrease), their safety rating also immediately increases/decreases by 0.4 to minimize bouncing between levels. This immediate increase/decrease generates the wealth kink we are examining in this study. License levels are identified by both a license level and a sublicense level (from zero to 4.99). In order to move to a higher license level, iRacers must complete a sufficient number of races with

[14] `https://www.esportsearnings.com/players/54552-maximilian-benecke`

an increasing number of corners per incident (CPI). As CPI increases, there is evidence of the iRacer's increasing skill in avoiding incidents, which suggests increasing likelihood of contributing to a safe racing environment. iRacing has a designated system by which iRacers can earn license levels, in which increases in safety score and license level are increasingly difficult to earn. This system is illustrated in Figure 3. The chart illustrates the non-linear relationship between corners per incident and safety score, and shows the kink points for each license level (the orange line is a Class D/beginner but not rookie — their safety score increases pretty quickly the cleaner their record. The black line is for a Pro. They have to drive cleanly for many more corners to improve. It also shows how the size of the kink is much larger as you move toward pro-level licenses.

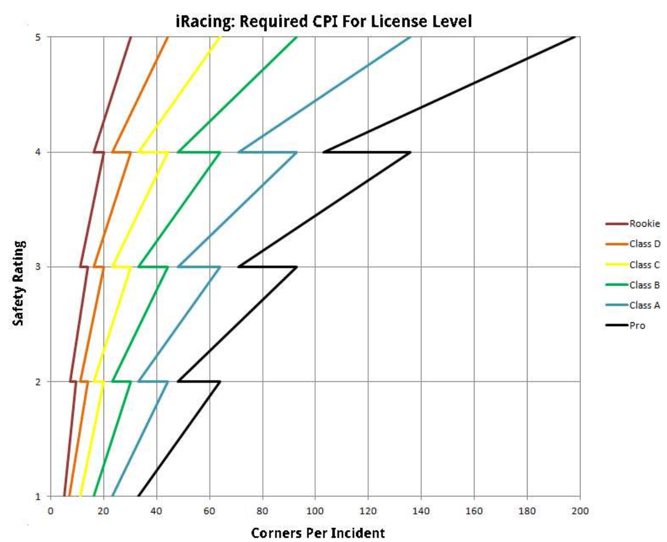

Figure 3: iRacing License Level Chart. Source: iRacing user guide.

While a driver's safety score may change with each race, license promotions occur only at the end of a 12-week season. Demotions also occur at the end of a season; however, if an iRacer's safety rating falls below 1.0, demotion happens immediately.

3.5 iRacing Series

Within each of the four racing categories, iRacing schedules dozens of official 12-week series. Each series runs with a particular car (e.g., Formula Renault 2.0) or class of car (e.g., GT3) for the entire season, changing tracks weekly. Races run as frequently as hourly or as infrequently as weekend only, depending on the popularity of the series. All series have a minimum license class required to participate. Rookie series all use the base iRacing content included with subscriptions and include lower-powered cars at less-challenging tracks. The fastest cars and most challenging tracks are limited to series requiring higher license levels (C, B, or A) For example, iRacers who want to compete in the popular *VRS GT3 Sprint* series must first earn and maintain a B license.

3.6 Race Process

Once an iRacer registers for a race, they must either participate or forfeit – the equivalent of having finished last in the race. iRacing divides the registered drivers into splits based on (1) the number of cars accommodated on the track's starting grid and/or pit lane and (2) the drivers' iRatings. If 60 iRacers register for a race on a track that accommodates 20 cars, the 20 drivers with the highest iRatings will comprise the top split, the 20 drivers with the lowest iRatings will comprise the bottom split, and the 20 remaining drivers will comprise the middle split. A brief practice session ensues, followed by a qualifying session, where drivers have a set number of minutes and/or laps to set their single best lap time. The fastest qualifying time among the drivers in the split will start the race on pole position, with each slower qualifying time on the grid behind that driver. Drivers who fail to qualify (either by choice or because they could not complete an incident-free qualifying lap) will start on the grid behind qualifiers, with their iRating determining their grid position. Races last a set number of minutes or laps, with penalties automatically assessed throughout the race for false starts, cutting corners, speeding in the pit lane, etc. Such penalties would require a driver to enter the pit lane for a stop-and-go or stop-and-hold for a number of seconds. Drivers with damaged vehicles may enter the pit lane for repairs. If the vehicle is too damaged to drive to the pit lane, they can request a "tow" to their pit stall, which incurs a time penalty before the repair ensues. After the race, iRacing adjusts each driver's iRating and safety rating according to their performance on the track and strength of field (average iRating of all iRacers in the split).

3.7 Violations and Protests

iRacing has a protest system to minimize abuse of the platform's no-fault incident system. After races, members can protest competitors for violation of the *iRacing Sporting Code*, a

37-page document that all members must agree to abide by when registering. Violations include intentional wrecking, retaliation, blocking (swerving on a straight to avoid being passed), and abusive behavior or language in voice or text chat during a session or offline on the forum. iRacing staff review each protest claim and can serve warnings or consequences, including feature restriction or suspension/loss of membership.

4 Data

Our data are taken from race results from ten consecutive 12-week racing seasons, from December 2018 to May 2020. To ensure stability of the data generating process, we chose the ten seasons of the *Skip Barber Race Series* that occurred during this time period. All races involve the same vehicle, the *Skip Barber Formula 2000*, a 150-horsepower 4-cylinder open-wheeled race car capable of top speeds reaching 135mph.[15] Further details and rationale for this dataset are explored below.

4.1 Data collection process

iRacers have access to several datapoints about all active iRacers and their races, from macro-level trends to the times and incidents that make up an individual race lap. Data from iRacing are ideal for investigating risk-taking behavior, given the quantity of variables available and the quantity of races.[16] These data are served to iRacers via active web pages. To facilitate data collection, we used a publicly-available Python library to retrieve data on members[17], cars, tracks, series, race results, incidents, and lap results. We then imported the resulting CSV files into SAS and converted time stamps and safety ratings into equivalent values better suited to regression analysis.[18] The final data set is completely unique; we are

[15]Details regarding this car are available at `http://iracing.com/cars/skip-barber-formula-2000/`.

[16]When an iRacer forces a disconnect due to poor racing performance or frustration, it is known as a "rage-quit". We define a rage-quit as a race tagged as a disconnect with at least one incident in the race.

[17]Pursuant to Institutional Review Board requirements, personally identifiable information, including names, has been purged from the dataset and retained in a secure data repository by one of the authors, with a linking variable. Personally identifiable information is not necessary to conduct the analysis here.

[18]The program library used to collect our data is posted at `http://github.com/jeysonmc/ir_webstats`.

not aware of any other academic studies employing data from iRacing.[19]

During the ten 12-week series spanning a time frame of December 2017 to May 2020, over one million races were recorded across all four disciplines in iRacing. Each series differs with respect to license requirements, skill level, incident limits, duration of races, and frequency of races. This particular series was chosen for a number of reasons: (1) The series uses a mix of free tracks and paid tracks. (2) The car is not included in iRacing's free content and must be purchased for $11.95, representing a financial commitment of the drivers that race in this series. (3) The series requires a D-level or better license, but is popular among all driver levels. In D-level races, drivers may "fast repair" their car one time in a race simply by stopping in their designated pit stall and receive an instantaneous repair. Subsequent repairs can take several minutes to repair in the pit stall, depending on the severity of the damage. This fast repair feature does not exist in C and higher level race series[20]. (4) Races are duration-consistent – around 25 minutes per race. (5) 99.44% of races "went official," meaning they had a minimum of six drivers and race results impacted drivers' iRatings and safety ratings. (6) Races occur every hour, providing us with race data for all parts of the day and from around the world.

4.2 Data description

Our original dataset includes information about each race and driver in the series over a time period of December 2017 through May 2020. Each race's course, car model, and strength of field is identified. Within each race, iRacers' race statistics are recorded, including pre- and post-race iRating and safety rating, the number of incidents, type of incidents, qualifying time, starting position, average lap time, fastest lap time, etc. The original dataset includes 945,482 individual race results over 56,084 races/splits from 37,242 individual iRacers. We removed unofficial races from the dataset; this filter reduced our sample size to 945,097 iRacer results over 55,985 races/splits and 37,241 iRacers. Since sim racing relies on constant internet connections and sufficient local processing power to maintain an iRacer's participation in the competition, we have deleted observations where only one lap was attempted before disconnecting and for which no incidents occurred. Since an iRacer who

[19] According to iRacing's Terms of Use, use of the data for purposes other than racing requires explicit written permission from iRacing. iRacing has generously granted permission to us to use the data for the purpose of academic research on risk-taking.

[20] All races in the Skip Barber series are D-level races, but higher license level drivers participate.

crashes in the first lap might be tempted to “rage-quit”, disconnects with one lap attempted and at least one incident were assumed to be intentional and were retained in the dataset. By deleting assumed connection-related (i.e., not “rage-quits”) disconnects, our final sample size falls to 806,253 individual race results over 55,985 races/split and 36,005 iRacers.

For unskilled participants, in may be difficult to disentangle incidents resulting from lack of skill and incidents resulting from risk-taking behavior. Thus, we omitted all racers who held the rookie license class, since iRacers are not assigned an iRating until they reach their D license. In addition, we deleted race results for which the driver’s result was “Disqualified/Scoring Invalidated”. These filters reduced our sample size to 764,972 race results over 55,985 races/splits and 34,565 iRacers. Finally, since racers learn strategy over repeated attempts, we deleted race outcomes for users with a limited number of races during the sample period.

Table 1: Descriptive Statistics

Variable	Mean	Std Dev	Minimum	Maximum
Race Incidents per Corner	0.4485	0.2193	0.0087	2.4737
iRating (000)	1.6230	0.9231	0.0280	8.2330
iRating Std. Dev. (by race)	0.4185	0.4207	0.0099	2.3684
Safety Score	1.8481	1.0290	-1.0000	5.9900
Week 12 indicator	0.0724	0.2591	0.0000	1.0000
Cusp Up Indicator	0.1001	0.3002	0.0000	1.0000
Cusp Up * Safety Score	0.1853	0.6405	-0.2000	5.9900
Cusp Up * Week 12	0.0061	0.0776	0.0000	1.0000
Cusp Down Indicator	0.0812	0.2732	0.0000	1.0000
Cusp Down * Safety Score	0.1216	0.4856	-0.9100	3.1000
Cusp Down * Week 12	0.0073	0.0849	0.0000	1.0000
Number of Cars Purchased	13.2941	10.7443	1.0000	67.0000
Size of Split Field	17.2159	2.1726	6.0000	20.0000
Track Paid Indicator	0.7464	0.4351	0.0000	1.0000
Observations	764,999			

4.3 Dependent variable

Our dependent variable in the analyses reported below is ”incidents per corner (IPC)” or the number of incident points a driver records in a race split, divided by the number of corners the driver achieved in that split. That is, a driver who crashes, disconnects, or fails to finish the race may have a different number of corners in the race than another driver who completes the race. This variable is used to approximate the risk-taking level by the driver. Recognizing that this is not a perfect correlation to risk-taking, since an incident can be recorded as a result of another driver's risk-taking, we still believe this ratio to be a good approximation. In the complete sample, drivers experience an average of 0.04 incident points per corner. When we eliminate racers without a sufficient number of splits in a series, the average incidents per corner falls slightly.

In Table 1, we list the independent variables used in our model, along with their interpretation and our intuition for why they might be explanatory. Our independent variables include driver level, split-level, and driver-split level variables. Driver level variables do not vary over the ten seasons, but driver-split level variables may vary for each iRacer in each race (split). Split-level variables are the same for each iRacer in a split. Descriptive statistics are given in Table 1.

4.4 Driver-level independent variables

Driver-level variables are collected once for all drivers after the conclusion of the series under investigation. While these variables may vary across the dataset, time-varying data were not available at the split-level at the time of collection. That said, we do not expect that any variation across these driver-level variables to be related to risk-taking behavior at a meaningful level.

iRacers with a larger financial investment in the game may have different risk preferences than those with smaller financial investment. While our data do not include all financial investments, such as hardware equipment, internet speed, and subscription discounts, we do know how many additional vehicles racers have purchased within the game.[21] On average, racers have 13.43 purchased vehicles in the game, and as many as 67. Relatively few racers have a very large inventory of vehicles; only 335 racers have 50 or more vehicles. The base cost of a car is $11.95, with increasing volume discounts for three, six, or 40 cars purchased.

[21]Due to volume discounts, the relationship between number of vehicles and financial investment may not be linear.

Because risk preferences may vary based on regional location, we employ regional fixed effects based on the racer's region of registration. Region size ranges from a single state in the U.S. to a group of small countries in the Middle East.

4.5 Split-level independent variables

There are split-level statistics that can impact both a driver's risky decisions and the overall probability of an incident. For each race session, iRacers are divided into splits based on license level and qualifying times. Some splits will have a tightly concentrated group of racers, where racers have similar iRatings. Other splits will have wider variation in skill level. Thus, we measure the standard deviation of iRating (divided by 1,000) within a split. A high standard deviation within a split indicates a wide range of racer skill, while a low standard deviation within a split indicates a high level of competitiveness. The average standard deviation within a split is 0.42.

Week 12 is an indicator variable taking the value of 1 if the split was run in the last week of the season (the "finale"). Drivers may shift their risk-taking strategy in the final week of the season, because license promotions and relegations usually happen at the end of the season. As a consequence, racers may shift their risk preferences and races may become more aggressive or tight in this last week.

Clearly, if a race has more incidents, the probability of a racer being involved in an incident increases. To control for split-level risk, we also include the total number of incident points recorded in a split divided by the number of corners in a split. On average, there were 0.0415 total incident points per corner at the split-level. Also, since more racers in a split increase the probability of an incident, we control for that increased risk by including the size of the field. The field size ranges from 6 to 20, with an average race field of 17.23.

At the split-level, we can also determine whether the track is a paid track or not. Paid tracks cost between $11.95 and $14.95, depending on the size and development effort involved. In a race on a paid track, all racers have additional financial investment in the game. 74% of driver-split results are on a paid track.

In addition to racer-split level variables and split-level variables, we also interact each cusp indicator with safety score and with the week 12 indicator. These four interactions allow us to explore the relationship between cusp and license level, as well as the relationship between cusp and final week effects.

4.6 Driver-split-level independent variables

The average iRating in our sample is 1,637 (divided by 1,000 for analysis purposes) and ranges from 28 to 8,729. While iRating is a zero-sum game, by dropping less experienced iRacers from our data, we have introduced some skewness into the iRating variable.

Safety score is a measure of the driver's license level at the beginning of the race. A driver with a low score has a safety record just above a rookie, while a driver with a high score may be a pro, or has raced cleanly for a long period of time. Within the game, license class is indicated by a letter (R[ookie], D, C, B, A, and P[ro]). This class is determined by a mapping of license level and license sublevel, with all promotions and most relegation happening at the end of the season. In the data, license level is reported as an integer between 5 and 28 and license sublevel is reported as an integer between 0 and 500. We combine these variables and convert them to a safety score, subtracting 4 from the license level and dividing by 4, retaining the integer. We then add the fraction that results from dividing the license sublevel by 500. Thus, our safety score matches the Safety Rating as shown in Figure 3. Consider, for example, a iRacer with a license level of 10 and a sublicence level of 282. To calculate their safety score, we follow this process: To calculate the integer of the safety score: $10 - 4 = 6$, $6 \backslash 4 = 1$, where $\backslash$ represents the integer division operator. To find the decimal portion of the safety score: $282/500 = 0.564$. Thus, this iRacer's safety score is 1.564.

Cusp Up is an indicator variable that is equal to 1 when the driver is close to achieving the next license level and 0 otherwise. As shown in Figure 3, once a driver moves to the next license level, their safety score will be increased by a fixed value to avoid frequent changes in license level. This discontinuity provides a kink that we will explore in the hypotheses. Cusp Down is, likewise, an indicator variable that is equal to 1 when the driver is close to relegation to the next lower license level and 0 otherwise. We use the license sublevel to determine the cusp, and the size of the cusp depends on the license class. A Class D driver is on the cusp if they are within 25 license sublevel points of a promotion or relegation. The cusp is 20 points for Class C drivers, 18 points for Class B drivers, 16 points for Class A drivers, 15 points for Pro Level 2 drivers and 10 points for Pro Level 1 drivers. The decreasing size of the

Table 2: Cusp Levels

License Level	Cusp Distance
Pro 1	10
Pro 2	15
A	16
B	18
C	20
D	25

cusp corresponds to the decreasing point shift around the license level promotion/relegation. Table 2 shows the cusp limits for each license level. Approximately 9% of the racer splits in the sample have the racer on the cusp of moving up, and approximately 7% of the racer splits have the racer on the cusp of moving down.

5 Hypotheses

Assuming that racers are expected iRating maximizers, their goal in the game is to achieve the highest iRating possible under all given rules and constraints. Since iRating is (essentially) a zero-sum game, this means that a racer's iRating gains come at the expense of other racers' outcomes (gains and losses). A racer can add iRating points by placing well in races, but can also lose points by placing poorly. Placing well, especially in competitive races, requires risk-taking and racers can invest in loss control by driving more conservatively. A racer can increase iRating faster by racing in more selective races, and access to those "splits" is achieved by higher license levels/safety scores. Safety scores are negatively impacted by incidents: to increase iRating more quickly, racers will need to limit the number of incidents.

5.1 Consistency Control Hypotheses: H1-H3

First, racers with high iRatings tend to be the most experienced, and are thus more likely to minimize unforced errors. As a result, we anticipate that racers with high iRatings experience fewer incidents per corner, giving rise to our first (alternative) hypothesis - always assuming that the same hypothesis in null form assumes independence:

Hypothesis 1 *Incidents per corner (IPC) is negatively related to iRating.*

This hypothesis constitutes a consistency control within our data, and serves as a basis for further analysis. Second, given that incidents arise from the actions of others, the more homogeneous a racer experience is within a split, the less likely a racer may be to suffer many incidents. Conversely, the more heterogeneous racers are within a split, the more likely a racer will be to experience many incidents. This leads us to our second consistency controlling hypothesis:

Hypothesis 2 *IPC is positively related to split-level standard deviation of iRating.*

Third, similar to the relationship between incidents and iRating, it should be expected that racers with a higher safety score tend to drive more skillfully and/or carefully, thereby reducing the propensity for incidents. Thus, we propose the next (alternative) hypothesis:

Hypothesis 3 *There is a negative relationship between IPC and safety score.*

It will be shown that the data consistently support these (alternative) hypotheses 1-3 (with p-values ¡1%). The next section will study reference-dependent preferences and introduce our first research questions.

5.2 Reference-dependent Preferences: H4-H6

The conceptual frameworks delineating reference-dependent preferences, characterized by a heightened sensitivity to losses in comparison to equivalent gains, trace their roots to the seminal prospect theory advanced by Kahneman and Tversky in 1979. Subsequently, these models have garnered substantial empirical support, elucidating puzzling phenomena such as the endowment effect. Their explanatory power extends to diverse domains, encompassing anomalies in labor market decisions, consumer behavior, and financial contexts, among others.

In the nascent stages of decision science, a discourse emerged highlighting incongruities in risk aversion observed across various utility elicitation methods. For instance, Sprenger

(2015) shows that, within the subset of expectation-centric models, a noteworthy prediction can be made - referred to as the "endowment effect for risk" - suggesting that alterations in reference points from certainty to stochastic conditions can induce variations in risk attitudes. Through two meticulously designed risk preference experiments, deliberately mitigating commonly discussed confounding factors, he empirically establishes the existence of an endowment effect for risk both across and within subjects. These findings not only contribute to the delineation of expectations-based reference-dependent models but also facilitate the assessment of recent theoretical extensions. Moreover, they offer valuable insights that may contribute to resolving the longstanding debate within decision science regarding inconsistencies in utility elicitation methodologies.

In simulated auto racing, racers consider reference points such as being likely to be promoted or delegated. We make a series of predictions related to behavior on the cusp of a promotion or relegation. An important prediction is that racers on the cusp of a promotion will tend to self-insure by taking less risk, a behavior likely to be demonstrated by a smaller number of incidents at the cusp of promotion because the racer will want to increase the probability of a large safety score jump. Conversely, racers at the risk of relegation will likely have a larger margin of error and will take more risk. Given a reference point, this behavior is consistent with Prospect Theory where agents exhibit risk aversion when facing potential gains (lower risk taking) and risk seeking when facing potential losses (higher risk taking). The prediction gives rise to a compound fourth (alternative) hypothesis as follows.

Hypothesis 4a *Racers on the cusp of promotion take less risk than racers not on the cusp, entailing fewer IPCs.*

Hypothesis 4b *Racers on the cusp of relegation take more risk than racers not on the cusp, entailing a higher number of IPCs.*

Building on the theory of Kőszegi and Rabin (2006) and Kőszegi and Rabin (2007), with respect to the interaction between cusp and safety score, we also hypothesize that higher-classed racers on the cusp of promotion will take *more* risk to ensure promotion than lower-classed racers; consistently, higher-classed racers on the cusp of relegation will take *less* risk to avoid potential relegation and *maintain their accomplished status quo* in the game (endowment effect):

Hypothesis 5a *Higher-classed racers on the cusp of promotion take more risk than racers not on the cusp.*

Hypothesis 5b *Higher-classed racers on the cusp of relegation take less risk than racers not on the cusp.*

The idea behind this effect is that, in particular in the case of H5b, a form of the "endowment effect for risk" kicks in: Individuals assign a high value to what they have already accomplished in the game, and as a consequence, due to higher (financial and emotional) investment, they tend to be more risk averse when on the cusp on relegation. It is noteworthy that, from a driver's perspective, delegation may be more likely but is still stochastic due to the behavior of other drivers. The endowment effect may also represent the tendency for drivers who own a license level to value it higher than individuals who do not own it.[22]

Another hypothesis focuses on the financial investment a racer has made in their game participation. While all racers pay a standard subscription fee (with volume and other discounts available), we do not observe their subscription details. Furthermore, we do not know what exact equipment they use to play the game (which can range from a keyboard and mouse to a full simulated race car rig, as shown in Figure 2). However, we can observe how many cars and/or tracks a racer has purchased. Due to volume discounts, and the opportunity to purchase "everything" for a discounted lump sum, this relationship is not linear; however, we attempt to estimate the relationship between financial investment and risk-taking using the number of add-ons purchased in the game, leading to

Hypothesis 6 *Racers with more paid cars and/or tracks take more risk.*

5.3 End-of-contract Stage Decision-Making: H7-H8

Our next hypotheses relate to week 12 risk-taking. Remember that this week is the most important within a 3-month period (a "season"). Since a racer's safety score at the end of

[22]See, for instance, Morewedge and Giblin (2015). Experimental evidence supports the existence of this investment effect for risk taking decisions. Knetsch and Sinden (1984) demonstrate that a higher fraction of individuals are willing to pay $2 to keep a lottery with unknown odds of winning around $50 than to accept $2 to give up the same lottery if they already own it. Kachelmeier and Shehata (1992) show that selling prices for a 50-50 gamble over $20 tend to be significantly higher than subsequent buying prices out of experimental earnings for the same gamble.

week 12 dictates their license level for the next 'season', we anticipate that behavior may shift from iRating maximization to safety score maximization. Do racers take less or more risk in week 12? Further complicating the prediction, drivers can earn independent recognition for their season-level performance, which is related to neither iRating nor license level. Some drivers may be highly motivated by season recognition, but we cannot distinguish among motivations in our dataset. Since we have no prediction, we will dispense with the alternative form and predict the null as follows.

Hypothesis 7 *There is higher risk-taking in week 12 relative to other weeks.*

We further analyze the impact of license level on cusp risk-taking and also the impact of week 12 on cusp risk-taking. Again, since it is difficult to disentangle active risk-taking from the riskiness of the more challenging track used in week 12, we will propose only the null hypothesis:

Hypothesis 8 *There is no relationship between cusp position and risk-taking in week 12.*

These hypotheses refer to the insight that reference dependent preferences imply *bunching* of performance at or around the reference points, "where it matters".

6 Model and Results

6.1 Econometric specification

We estimate the relationships using the following standard OLS model:

$$IPC_{i,n} = \alpha + \beta_1 OldiRating_{i,n} + \beta_2 OldSafety_{i,n} + \beta_3 cusp_{i,n} + Y_{i,n}\beta + X_{i,n}\gamma \qquad (1)$$

where Incidents per Corner (IPC) is the dependent variable, $Y_{i,n}$ are interactions with the cusp indicator and $X_{i,n}$ are racer, split, and racer-split controls. We control for regional fixed effects with an indicator variable for each region represented in the dataset.[23]

[23] In a few instances, we have combined some small regions with other demographically similar regions.

Our initial results do not include cusp interactions and are summarized in Table 3. Each column of results represents a successively smaller sample, with columns indicating results from a sample that eliminates racers with fewer races in the series. Column 1 includes all qualifying racers as described earlier. Column 2 removes racers with fewer than five races in the series. Column 3 removes racers with fewer than ten races in the series. Finally, column 4 removes racers with fewer than 20 races in the series. Each specification yields qualitatively similar outcomes, speaking for consistency of our results.

6.2 Reference-Dependent Preferences: H1-H6

Hypothesis 1 is supported by the data: The results indicate that the iRating prior to a race is negatively related to the number of incidents per corner. Furthermore, Hypothesis 2 is also supported by the positive and significant coefficient on iRating standard deviation at the split-level. Hypothesis 3 is supported, as well, showing a negative and significant coefficient on safety score. In other words, as expected, racers with high safety scores experience fewer incidents per corner.

Of most interest for this study is the relationship between incident experience and the cusp. The negative and significant coefficient on Cusp Up corresponds with the positive and significant coefficient on Cusp Down, jointly affirming Hypotheses 4a and 4b. The remaining variables shown in **Table 3** support the corresponding Hypothesis with regard to financial investment in the game. The Size of Split variable is a control, but follows intuition in that the larger the split field, the more incidents racers experience.

Turning now to **Table 4**, we have added the interactions with the cusp indicators to test hypotheses 5 and 6. All other variable are consistent in both effect and significance when we add these interactions. The interactions of Cusp Up and Cusp Down with safety score yield the predicted results and support alternative hypotheses 5a and 6a. Racers on the cusp of promotion take more risk when they hold higher safety scores, and thus would enjoy a higher license level. Racers on the cusp of relegation take less risk, presumably to reduce the risk of relegation. This is consistent with an aspiration level where racers want to win at least something or at least not fall below a certain subsistence level - the aspiration level (Diecidue and Van De Ven (2008)). Notably, when we restrict the sample to racers with at least 20 races in the series, the interaction between Cusp Up and Safety Score is not significant any more.

Finally, the positive and significant relationship between financial investment in the racing game and IPC supports Hypothesis 6: A higher individual investment is associated with more IPC, ceteris paribus. An underlying reason might be because a driver has a higher incentive to improve their iRating in the game.

6.3 Findings relating to final-contract-stage behavioral patterns: H7-H8

Regarding the sign on the week 12 indicator (hypothesis 6), we find a positive and significant relationship, suggesting more aggressive racing in week 12, the final contract stage. This is interesting as it suggests more risk taking "when it matters", at least for the season recognition - an incentive that seems to overcompensate the effect of attaining a better license level in the short run. In accordance with Frick (2011) who use the German Bundesliga for soccer player's effort levels, we can at least say that when the outcome of individual effort is uncertain and risks are involved, racers tend to exert more effort to win and thereby increase their risk-taking at the margin, aiming for potential better status in final contract stages.

This effect is, however, not confirmed when a racer is also near a cusp or reference point. In other words, racers do not also adjust their effort at a certain point in time (for instance, the end of a quarter/season) based on their proximity to a reference point (Hypothesis 7).

Finally, we see from the interactions with week 12 that we cannot reject the null Hypothesis 8 for any specification. There is no evidence that we can disentangle incentive effects from track difficulty in week 12. Future research can attempt to identify other factors that might reveal a meaningful interaction with week 12.

7 Concluding Remarks

This paper utilizes data from over 56,000 simulated auto races globally to examine risk-taking behaviors at the margins, aligning with reference-dependent preferences. The study reveals that participants' risk-taking behavior changes when a desired intermittent outcome is introduced, sometimes at the expense of achieving a more favorable expected end state. Specifically, we find that intermediate kinks in the utility function lead racers to take less risk when opportunities to increase temporal wealth arise, and more risk when there is a potential to lose temporal wealth. This provides important insights into the incentives for risk-taking at the margins of wealth kinks (e.g., retirement age, family changes, etc.). Risk aversion intensifies at these kinks, while risk-taking escalates with greater financial investment.

Individual incentives are notably linked to the license level available: Racers on the verge of attaining higher license levels tend to take more risk, whereas those at risk of being relegated to a lower license level exhibit less risk-taking due to the fear of losing their current status for the next season. Furthermore, risk-taking increases with the extent of actual financial investment in the game. Risk aversion appears to strengthen at kinks in the payoff curve.

Our study reveals findings consistent with existing literature on the impact of final contract stages on individual decision-making. Previous economic studies in sectors like sports and public procurement have noted increased individual effort during these critical contract stages. For instance, a football or soccer player nearing the end of their contract may exert more effort to demonstrate their value, while a coach, after securing a long-term contract, may reduce effort initially. In scenarios where individual effort outcomes are uncertain and involve risks, it makes sense for individuals to increase risk-taking when aiming for a potential better status and reduce it in other situations to protect their current status. Our study evaluates changes in individual risk-taking as participants approach situations where performance becomes more critical. It contributes to the literature on risk-taking by providing insights into behavioral adjustments during final contract stages related to individual risk-taking at the margins.

These findings offer significant insights into the incentives for risk-taking at the margins of wealth kinks. They provide a foundation for analyzing other wealth kinks, including contract negotiations, age-related breaks in investment planning, and financial decisions influenced by changes in family size due to birth, death, or the maturation of children.

Table 3: Empirical Results [Dependent Variable: IPC]

Parameter	All racers		At least 5 splits		At least 10 splits		At least 20 splits		At least 50 splits	
Intercept	0.0211	***	0.0205	***	0.0201	***	0.0203	***	0.0208	***
	(0.0009)		(0.0009)		(0.0009)		(0.0010)		(0.0012)	
iRating (000)	-0.0028	***	-0.0028	***	-0.0027	***	-0.0026	***	-0.0023	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0002)	
iRating Std. Dev. (by race)	0.0016	***	0.0015	***	0.0014	***	0.0014	***	0.0009	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0003)	
Safety Score	-0.0072	***	-0.0072	***	-0.0072	***	-0.0073	***	-0.0076	***
	(0.0001)		(0.0001)		(0.0001)		(0.0002)		(0.0002)	
Week 12 indicator	0.0039	***	0.0038	***	0.0038	***	0.0039	***	0.0035	***
	(0.0003)		(0.0003)		(0.0003)		(0.0003)		(0.0003)	
Cusp Up Indicator	-0.0036	***	-0.0034	***	-0.0032	***	-0.0027	***	-0.0023	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0003)	
Cusp Down Indicator	0.0028	***	0.0028	***	0.0025	***	0.0022	***	0.0018	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0003)	
Race Incidents per Corner	0.0500	***	0.0495	***	0.0487	***	0.0481	***	0.0471	***
	(0.0004)		(0.0004)		(0.0004)		(0.0005)		(0.0006)	
Number of Cars Purchased	0.0001	***	0.0001	***	0.0001	***	0.0001	***	0.0001	***
	(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)	
Size of Split Field	0.0005	***	0.0005	***	0.0005	***	0.0005	***	0.0005	***
	(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)	
Track Paid Indicator	0.0044	***	0.0045	***	0.0045	***	0.0044	***	0.0044	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0002)	
N	792,665		764,999		718,791		634,106		460,910	
r-Square	9.423%		9.402%		9.408%		9.630%		9.938%	

Standard errors (clustered at the driver level) in parentheses.=p-value < 10%; **=p-value < 5%; ***=p-value < 1%

Table 4: Empirical Results [Dependent Variable: IPC]

Parameter	All racers		At least 5 splits		At least 10 splits		At least 20 splits		At least 50 splits	
Intercept	0.0210	***	0.0205	***	0.0200	***	0.0202	***	0.0207	***
	(0.0009)		(0.0009)		(0.0009)		(0.001)		(0.0012)	
iRating (000)	-0.0028	***	-0.0028	***	-0.0027	***	-0.0026	***	-0.0023	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0002)	
iRating Std. Dev. (by race)	0.0015	***	0.0015	***	0.0014	***	0.0014	***	0.0009	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0003)	
Safety Score	-0.0072	***	-0.0071	***	-0.0072	***	-0.0073	***	-0.0075	***
	(0.0001)		(0.0001)		(0.0001)		(0.0002)		(0.0002)	
Week 12 indicator	0.0039	***	0.0038	***	0.0038	***	0.0038	***	0.0034	***
	(0.0002)		(0.0003)		(0.0003)		(0.0003)		(0.0003)	
Cusp Up Indicator	-0.0071	***	-0.0068	***	-0.0064	***	-0.0056	***	-0.0050	***
	(0.0004)		(0.0004)		(0.0005)		(0.0005)		(0.0007)	
Cusp Up * Safety Score	0.0019	***	0.0018	***	0.0017	***	0.0015	***	0.0013	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0003)	
Cusp Up * Week 12	0.0006		-0.0003		-0.0003		-0.0007		-0.0007	
	-0.0009		(0.0008)		(0.0008)		(0.0008)		(0.001)	
Cusp Down Indicator	0.0080	***	0.0077	***	0.0072	***	0.0067	***	0.0060	***
	(0.0005)		(0.0006)		(0.0006)		(0.0006)		(0.0007)	
Cusp Down * Safety Score	-0.0035	***	-0.0033	***	-0.0031	***	-0.0030	***	-0.0027	***
	-0.0002		(0.0003)		(0.0003)		(0.0003)		(0.0003)	
Cusp Down * Week 12	-0.0006		-0.0004		-0.0001		0.0004		0.0011	
	(0.0009)		(0.0009)		(0.0009)		(0.0009)		(0.0011)	
Race Incidents per Corner	0.0500	***	0.0494	***	0.0487	***	0.0481	***	0.0471	***
	(0.0004)		(0.0004)		(0.0004)		(0.0005)		(0.0006)	
Number of Cars Purchased	0.0001	***	0.0001	***	0.0001	***	0.0001	***	0.0001	***
	(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)	
Size of Split Field	0.0005	***	0.0005	***	0.0005	***	0.0005	***	0.0005	***
	(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)		(<0.0001)	
Track Paid Indicator	0.0044	***	0.0044	***	0.0045	***	0.0044	***	0.0044	***
	(0.0002)		(0.0002)		(0.0002)		(0.0002)		(0.0002)	
N	792,665		764,999		718,791		634,106		460,910	
r-Square	9.474%		9.449%		9.451%		9.667%		9.971%	

Standard errors (clustered at the driver level) in parentheses.= p-value < 10%; **=p-value < 5%; ***=p-value < 1%

References

Albright, A. (2018). Cowboys. Rangers. Fortnite? A Texas city bets on eSports stadium. bloomberg.com/news/articles/2018-08-31/cowboys- rangers-fortnite-texas-city-bets-on-esports-stadium. Last accessed: May 24, 2019.

Allais, M. (1953a). Fondements d'une théorie positive des choix comportant un risque et critique des postulats et axiomes de l'école américaine. *Colloques Internationaux du Centre National de la Recherche Scientifique (Econometrie) 40*, pages 127–140.

Allais, M. (1953b). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*, pages 503–546.

Allen, E. J., Dechow, P. M., Pope, D. G., and Wu, G. (2017). Reference-dependent preferences: Evidence from marathon runners. *Management Science*, 63(6):1657–1672.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of economic perspectives*, 27(1):173–196.

Brocas, I., Carrillo, J. D., Giga, A., and Zapatero, F. (2019). Risk aversion in a dynamic asset allocation experiment. *Journal of Financial and Quantitative Analysis*, 54(5):2209–2232.

Brown, J. (2008). iRacing.com racing simulator: The answer to all your racing prayers — and your budget. caranddriver.com/features/iracingcom-racing-simulator/. Last accessed: May 24, 2019.

Brunnermeier, M. K. and Nagel, S. (2008). Do wealth fluctuations generate time-varying risk aversion? micro-evidence on individuals. *American Economic Review*, 98(3):713–36.

Chan, S. (2018). Mobile esports ascends in Asia. venturebeat.com/2018/02/08/newzoo-mobile-esports- ascends-in-asia/. Last accessed: June 6, 2019.

Coulom, R. (2007). Computing Elo ratings of move patterns in the game of Go. *Icga Journal*, 30(4):198–208.

Diecidue, E. and Van De Ven, J. (2008). Aspiration level, probability of success and failure, and expected utility. *International Economic Review*, 49(2):683–700.

Dionne, G. and Li, J. (2011). The impact of prudence on optimal prevention revisited. *Economics Letters*, 113(2):147–149.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J., and Wagner, G. G. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3):522–550.

Eeckhoudt, L. and Gollier, C. (2005). The impact of prudence on optimal prevention. *Economic Theory*, 26(4):989–994.

Frick, B. (2011). Performance, salaries, and contract length: Empirical evidence from german soccer. *International Journal of Sport Finance*, 6(2):87.

Fudenberg, D. and Tirole, J. (1990). Moral hazard and renegotiation in agency contracts. *Econometrica: Journal of the Econometric Society*, pages 1279–1319.

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

Greer, W. R. (1974). Theory versus practice in risk analysis: An empirical study. *Accounting Review*, pages 496–505.

Greer, W. R. and Skekel, T. D. (1975). Theory versus practice in risk analysis: A reply. *Accounting Review*, pages 839–843.

Halek, M. and Eisenhauer, J. G. (2001). Demography of risk aversion. *Journal of Risk and Insurance*, pages 1–24.

Hersch, P. L. and McDougall, G. S. (1997). Decision making under uncertainty when the stakes are high: Evidence from a lottery game show. *Southern Economic Journal*, pages 75–84.

Hoskins, C. G. (1975). Theory versus practice in risk analysis: A comment. *Accounting Review*, pages 835–838.

Ingraham, C. (2018). The massive popularity of esports, in charts. washingtonpost.com/business/2018/08/27/massive-popularity-esports-charts/. Last accessed: May 24, 2019.

Iossa, E. and Rey, P. (2014). Building reputation for contract renewal: implications for performance dynamics and contract duration. *Journal of the European Economic Association*, 12(3):549–574.

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

Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 121(4):1133–1165.

Kőszegi, B. and Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, 97(4):1047–1073.

Lawrence, A. (2018). How iRacing is democratizing motorsports. *The Atlantic*.

Morewedge, C. K. and Giblin, C. E. (2015). Explanations of the endowment effect: an integrative review. *Trends in cognitive sciences*, 19(6):339–348.

O'Neill, H. M. and Deacle, S. (2019). All out, all the time? evidence of dynamic effort in major league baseball. *Applied Economics*, 51(38):4191–4202.

Paravisini, D., Rappoport, V., and Ravina, E. (2017). Risk aversion and wealth: Evidence from person-to-person lending portfolios. *Management Science*, 63(2):279–297.

Peter, R. (2021). Who should exert more effort? risk aversion, downside risk aversion and optimal prevention. *Economic Theory*, pages 1259–1281.

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.

Rabin, M. and Thaler, R. H. (2001). Anomalies: Risk aversion. *The Journal of Economic Perspectives*, 15(1):219–232.

Riley, William B., J. and Chow, K. V. (1992). Asset allocation and individual risk aversion. *Financial Analysts Journal*, 48(6):32–37.

Russ, H. (2019). Global eSports revenues to top $1 billion in 2019: report. reuters.com/article/us-videogames-outlook-idUSKCN1Q11XY. Last accessed: May 24, 2019.

Samuelson, P. (1963). Risk and uncertainty: A fallacy of large numbers. *scientia*, 57(98).

Samuelson, W. and Zeckhauser, R. (1988). Status quo bias in decision making.

Segal, U. and Spivak, A. (1990). First order versus second order risk aversion. *Journal of Economic Theory*, 51(1):111–125.

Sinn, H.-W. (1982). Kinked utility and the demand for human wealth and liability insurance. *European Economic Review*, 17(2):149–162.

Sprenger, C. (2015). An endowment effect for risk: Experimental tests of stochastic reference points. *Journal of Political Economy*, 123(6):1456–1499.

Stiroh, K. J. (2007). Playing for keeps: Pay and performance in the NBA. *Economic Inquiry*, 45(1):145–161.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, 12(3):183–206.

Wakker, P. and Tversky, A. (1993). An axiomatization of cumulative prospect theory. *Journal of risk and uncertainty*, 7:147–175.