ECON30580 Economics of Betting Markets
12. Field Size and Margins Across Markets

Karl Whelan

School of Economics, UCD

Spring 2026

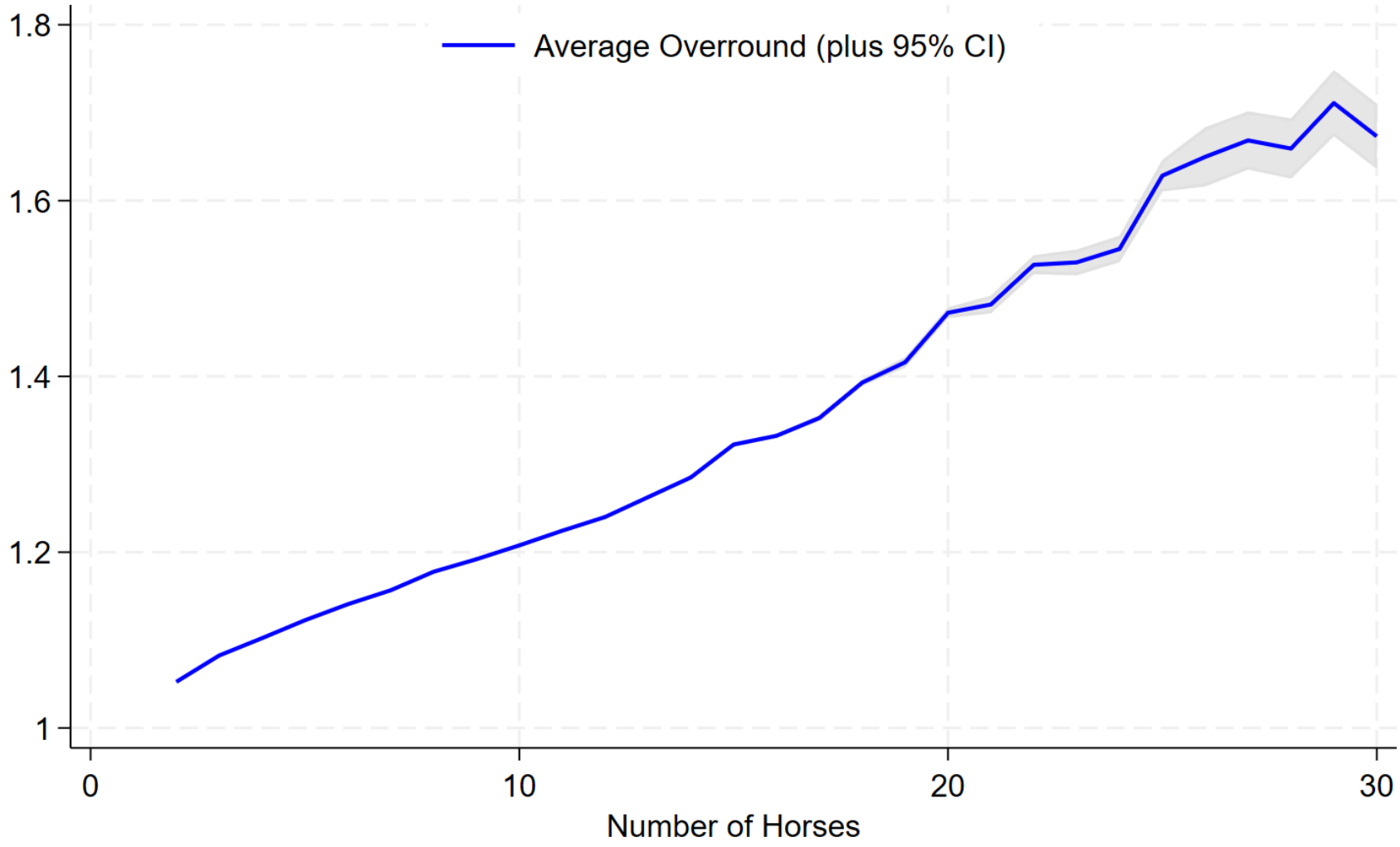
Part I

Field Size in Horse Racing

Overrounds Are Higher for Races with More Horses

- I downloaded a dataset with starting price odds for 396,572 horse races involving 4.1 million individual horses as well as data on the results.
- There is also an interesting pattern observed in horse racing: Overrounds are larger for races with more horses, meaning bookmakers earn higher margins on races with more competitors.
- For each race, I calculated the overround and charted it against the number of horses competing.
- See the graph two pages down.
- Average overrounds go from 105% for (literal) two-horse races to 167% for races with 30 horses.
- This pattern is well-known in the industry. Indeed, I have seen materials from horse racing bettor associations rating the odds offered by various tracks and operators by calculating the overround per horse, suggesting it is seen as a natural thing for overrounds to get worse as the number of horses increases.

Average overrounds sorted by number of horses in the race



Are Insiders the Explanation?

- One explanation for the increase in overround as the number of horses in the race increases is a theory we have heard before: Insiders.
- In one of his papers on the impact of insiders, Shin (1993) presented a model in which the overround would rise as the number of insiders increased.
- The idea is that as the number of horses goes up and there are more horses with high odds, then the bookmaker is more at risk because the insiders may know that a specific horse with high odds is going to win. To counter this risk, the bookmakers set worse odds on all the horses, leading to a high overround.
- You probably know already I'm not a big fan of insider theories, though even I would concede that insider knowledge probably matters more in horse racing than other sports.
- But evidence from our racing dataset points against the insiders explanation.
- A regression of overrounds on indicators that the race features young horses that aren't well known, so there is more chance for inside information (novice races, maiden races, races for 2-year olds and 3-year olds) doesn't explain much of the variation in overrounds (R^2 of only 0.0336) and three of the four coefficients are the "wrong" sign.

Regression of the overround on indicators the race features less well-known horses

```
.
. reg overround novice maiden two_yo three_yo
```

Source	SS	df	MS		
Model	232.754068	4	58.188517	Number of obs	= 377,855
Residual	6699.40928	377,850	.017730341	F(4, 377850)	= 3281.86
				Prob > F	= 0.0000
				R-squared	= 0.0336
				Adj R-squared	= 0.0336
Total	6932.16334	377,854	.018346143	Root MSE	= .13316

overround	Coefficient	Std. err.	t	P>|t|	[95% conf.	interval]
novice	-.0499611	.0007042	-70.95	0.000	-.0513413	-.0485809
maiden	.0470337	.0006269	75.02	0.000	.0458049	.0482625
two_yo	-.041221	.000745	-55.33	0.000	-.0426811	-.0397609
three_yo	-.0207233	.0004827	-42.93	0.000	-.0216695	-.0197772
_cons	1.233793	.0003462	3563.92	0.000	1.233115	1.234472

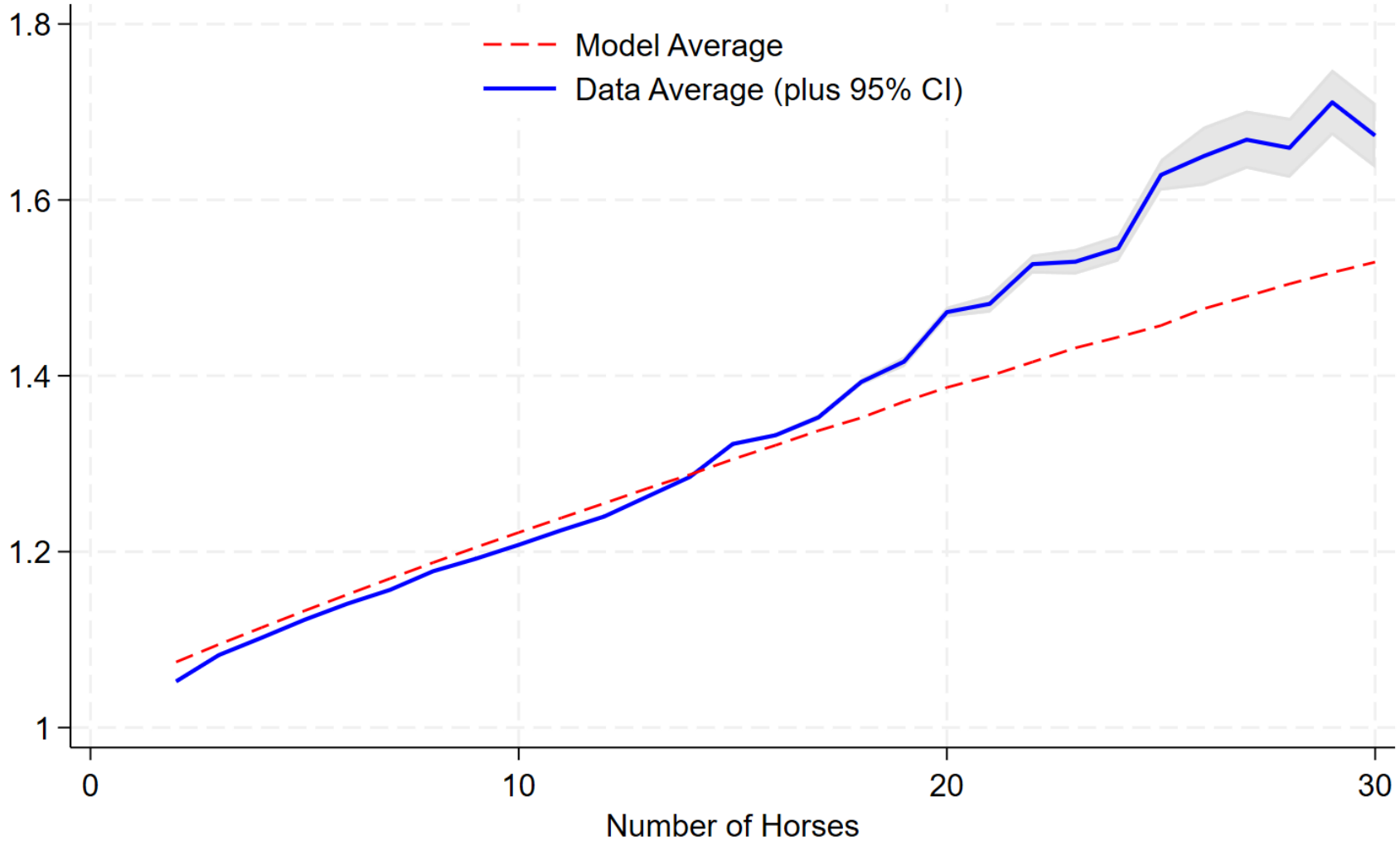
Could the Monopoly with Disagreement Model Explain It?

- Yes! The monopoly model predicts that the decimal odds for a horse with probability p of winning where beliefs about the probability are uniform on $[p-\sigma, p+\sigma]$ would be

$$D = \sqrt{\frac{1-\mu}{p\,(p+\sigma)}}$$

  where μ is the cost per bet of running the business.
- Given a set of probabilities and assumed values of σ and μ, we can calculate the odds and thus the overround.
- Using the normalised probabilities from the horse racing data set and again using $\sigma = 0.06$ and $\mu = 0.02$, the model's average overround by number of horses matches the data pretty well. See the next page.
- Why? Check the average normalised probabilities in races sorted by number of horses two pages down: In races with lots of horses, they are all longshots and so the bookmaker can offer them bad odds on all of them.
- For more than 18 horses, the model under-predicts the overround a bit. It may be that once there are loads of horses, you get into Grand National "bettors without a clue" territory. Or maybe it's insiders ...

Monopoly model's predictions for the overround versus the data



Average normalised probabilities for races with differing numbers of horses

Average normalised probabilities of the top 10 horses in races

	Number of Horses in the Race				
Horse Rank by Probability	*3*	*5*	*10*	*15*	*25*
1	56%	43%	29%	22%	17%
2	30%	25%	18%	14%	11%
3	14%	16%	13%	11%	8%
4		10%	10%	9%	6%
5		5%	8%	7%	6%
6			7%	6%	5%
7			5%	5%	4%
8			4%	5%	4%
9			3%	4%	4%
10			2%	4%	3%

Part II

Field Size and Margins in Other Sports

A Dataset of "Outright" Bets

- 137,402 price quotes on outright bets from *SportsOddsHistory.com*
- Examples: Super Bowl winner, Champions League winner, Wimbledon winner
- Prices collected across different points in time within a season or tournament
- 4,959 implied bookmaker margins calculated

Two Key Features

- **Large variation in field sizes**
  - NCAA tournament futures: $\sim$300 teams quoted each September
  - Golf tournaments: 100+ players at the start
  - Super Bowl: 32 teams $\rightarrow$ 14 playoff teams $\rightarrow$ 2 finalists
- **Little scope for private information**
  - Hard to have inside information on season-long outcomes
  - Unlikely bookmakers adjust odds broadly to defend against insiders

Implication

- If margins rise with field size here, the *inside information explanation fails*
- Data show margins **increase with field size**, consistent with the **rational disagreement model**

The Full Dataset

Quotes on outright bets in the dataset

Sport	*Years*	*Quotes*	*Margins*
MLB (World Series, AL, NL)	2005-2024	9,177	604
NBA (Championship, West, East)	2009/10 to 2024/25	10,172	635
NFL (Super Bowl, AFC, NFC)	2008/09-2024/25	17,183	1,010
NHL (Championship, West, East)	2009/10 to 2024/25	7,917	492
NCAA College Basketball	2009/10 to 2024/25	17,412	190
NCAA College Football	2000-2024	13,892	298
Football Leagues (England, Spain, Italy)	2009/10 to 2024/25	5,111	373
UEFA Champions League	2009/10 to 2024/25	1,539	92
Tennis Grand Slams	2000-2025	24,915	970
Golf Majors	2000-2025	30,084	295

Actual Margins versus Model Predicted

Average implied margins and the profit-maximising bookmaker model's predictions

Number of Competitors	*Data*	*Model Implied Margin*
2-3	4.2%	3.5%
4-6	9.0%	5.8%
7-10	13.5%	8.8%
11-14	17.4%	12.7%
15	16.8%	14.6%
16	18.3%	14.8%
17-26	17.6%	13.7%
27-31	21.6%	22.1%
32-64	23.6%	22.4%
Over 64	34.5%	34.7%

Margins fall over the course of seasons

Average implied margins for bets on who will win the World Series and NBA Championship at different points in the season

World Series Winner		*NBA Champions*	
December	22.4%	August	20.3%
April	24.4%	October	18.0%
May	20.8%	December	17.6%
June	19.7%	January	18.0%
All-Star Break	18.9%	February	17.2%
August	19.0%	March	17.1%
September	17.2%	April	15.7%
Wild Card	12.8%	Round 1	14.5%
Divisional Series	12.1%	Conference Semis	11.0%
League Championship	5.5%	Conference Finals	5.9%
World Series	3.5%	Finals	3.5%

Margins fall over the course of seasons

Average implied margins for bets on who will win the Super Bowl and Champions League at different points in the season

Super Bowl Winner		*Champions League Winner*	
Week 1	20.8%	June	24.3%
Week 8	20.7%	August	22.7%
Week 12	20.3%	October	20.6%
Week 17	17.0%	Knockout Playoff	20.3%
Wild Card	13.6%	Round of 16	14.5%
Divisional Round	11.0%	Quarter Finals	10.7%
Conference Finals	5.9%	Semi Finals	8.0%
Super Bowl	3.7%	Final	5.4%

Outright Bets and Possible Opportunities

- We have seen that outright bets (e.g. season winners) typically have large margins and deliver big average losses.
- One caveat: these markets are often **low priority** for bookmakers.
  - They are not very active day-to-day.
  - Odds may not be updated as frequently.
- This means it is sometimes possible to find **stale, mispriced odds**.
- However, the evidence suggests such opportunities are likely **rare**.

A practical alternative

- If you think a team will outperform expectations early in the season, consider related markets with **lower margins**.
- Examples:
  - "Will this team get 80 points or more?"
  - "Will this team finish above 6th?"
- These will typically be priced closer to 50-50 bets with low margins.

Part III

Margins Across Markets

Evidence for Soccer

- The next pages shows average margins (calculated from overrounds) for soccer betting on home, away or draw outcomes. for 22 European soccer leagues again using our dataset of 151,254 matches over the period 2005-2025.
- Note that margins are lowest for high profit leagues.
- When there are multiple leagues from the same country, the overrounds get bigger as we go from the top division downwards to the lower leagues.
- One explanation consistent with our pricing formula

$$D = \sqrt{\frac{1-\mu}{p\left(p+\sigma\right)}}$$

  is that the per-bet cost parameter μ might depend on volumes.
- Perhaps the total cost of running a market is actually of the form $F + cV$ where F is a fixed cost, c is a variable cost and V is volume. In that case, the unit cost parameter would fall with volumes, so high-volume events had better odds.

$$\mu = c + \frac{F}{V}$$

Average margins implied by overrounds

Average implied margins for 22 European leagues

League	*Implied Margins*
English Premier League	5.3%
Germany Bundesliga	5.7%
Spain La Liga	5.7%
Italy Serie A	5.9%
France Ligue 1	6.3%
English Championship	6.4%
Netherlands Eredivisie	7.0%
Scotland Premier League	7.0%
Portugal Primeira Liga	7.1%
Germany Bundesliga 2	7.1%
English League 1	7.1%
Turkey Super Lig	7.2%
English League 2	7.3%
Belgian First Division	7.4%
France Ligue 2	7.6%
Italy Serie B	7.8%
Spain La Liga 2	7.8%
Greece Super League	8.0%
English Conference	8.3%
Scotland Championship	8.5%
Scotland League 1	8.8%
Scotland League 2	9.0%

Volumes from Betfair

- Bookmakers don't tell us how many bets they took on each game.
- An alternative is to use information from **Betfair Exchange**, a completely different type of betting market.
- On exchanges, the platform arranges peer-to-peer matching and ordinary bettors can agree to take the opposite sides of a bet.
- We will discuss betting exchanges later. What matters now is that Betfair makes the volumes traded on its exchange available publicly.
- We can use these as a proxy for the relative volumes likely placed with bookmakers.
- The next page shows average pre-match volumes from 2022-24 on Betfair Exchange for the 22 leagues.
- Clearly, the volumes line up well with margins. Confirmation that high volumes contribute to lower margins.

Average volumes

Average volume before kick-off on Betfair for 22 European leagues

League	*Average Volume*
English Premier League	£914,063
Italy Serie A	£417,729
Spain La Liga	£371,811
Germany Bundesliga	£247,742
France Ligue 1	£141,250
English Championship	£94,299
Turkey Super Lig	£70,733
Portugal Primeira Liga	£69,070
Scotland Premier League	£57,025
Netherlands Eredivisie	£46,734
Germany Bundesliga 2	£42,438
Belgian First Division	£30,283
Spain La Liga 2	£28,542
English League 1	£27,618
Italy Serie B	£24,810
Greece Super League	£23,700
English League 2	£18,274
France Ligue 2	£12,006
Scotland Championship	£6,025
English Conference	£5,819
Scotland League 1	£4,550
Scotland League 2	£3,228

Other Explanations: Variations in Disagreement

- Matching up Betfair volumes for 2022-24 with our dataset of European soccer margins, the volumes are highly significant as a predictor.
- But a model with just volumes has a much lower R^2 than a model with league dummies, so volumes do not fully explain why the high-profile leagues have so much lower margins.
- Another explanation? Maybe there is more disagreement about matches in lower profile leagues so σ is higher.
- One way to measure this is in the cross-sectional variance of odds offered by bookmakers on matches. www.football-data.co.uk includes individual bookie's quotes as well as averages.
- This also adds some explanatory value: The R^2 is improved by a few points by adding the cross-sectional variance in odds.
- But the fit is still a long way below what league effects on their own explain.
- Something else is going on.

Risk in the Bookmaking Business

- The Law of Large Numbers helps make bookmaking less risky than it might appear: Averaging over many events, bookmakers will expect a bet with a probability p of winning to win a fraction p over time.
- Bookmakers may lose money on a particular match, but over many bets they expect to come out ahead on average.
- However, they do not ignore risk altogether.
- One important risk is that their **odds may simply be wrong**.
  - A bettor with a better model could repeatedly pick them off.
- For major sports such as the Premier League or the NBA, bookmakers have large datasets and considerable experience.
- For lower-profile sports, they are far less confident about their probability estimates.

Ambiguity Aversion

- In practice, probabilities are rarely things we truly know.
- They are just our **best estimates**.
- Standard economic decision theory says this is fine:
  - Use your best estimate of the probability and make the optimal choice.
- But experiments suggest people dislike situations where probabilities are themselves uncertain, a phenomenon known as **ambiguity**.

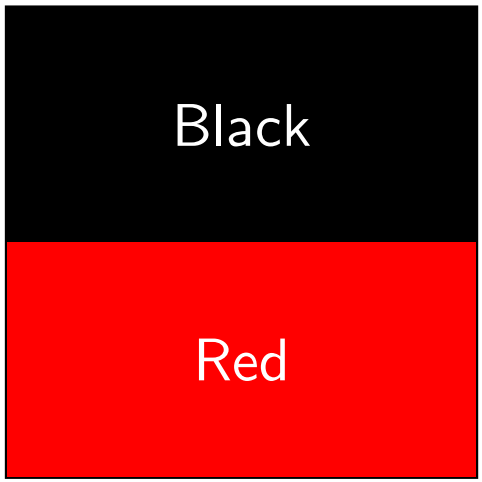
Ellsberg Paradox

- Box A: Equal numbers of red and black balls.
- Box B: Only red and black balls, but the mix is unknown.
- You win a prize if you draw a red ball.
- Rationally, you should treat the unknown box as a 50â50 chance.
- In practice, most people prefer the **known probabilities**.

Conclusion: People dislike ambiguity as well as risk.

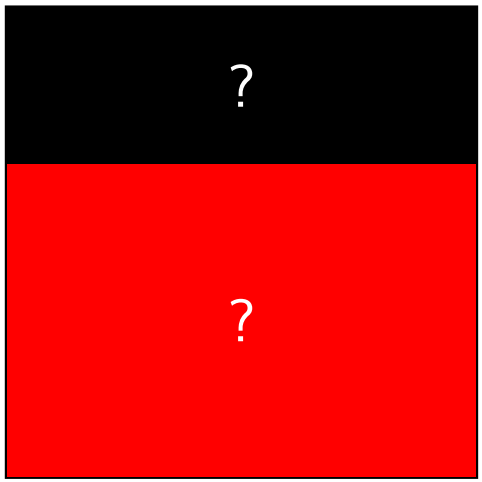
The Ellsberg Paradox

Box A (Known Probabilities)



- Equal numbers of red and black balls
- Win if you draw a red ball
- Probability = 50%

Box B (Unknown Probabilities)



- Only red and black balls
- The mix is unknown
- Win if you draw a red ball

Ambiguity and Bookmaker Margins

- Bookmakers face the same problem.
- They feel confident about probabilities for major leagues.
- But they are far less certain about **obscure markets**.
- These markets feel **riskier** because the probabilities are harder to estimate.

Marco Blume (Pinnacle)

> *"In the smaller leagues, we have zero information but so does everybody else. I mean, who really knows what's happening in Thailand's second division? We try to almost keep things on a level playing field by adjusting the margins ..... While we would take a tremendous amount of wagers on a La Liga game or NBA game or NFL game with a very low margin ... if you go down the tier leagues, the overround is increasing and the limits are reducing. So, it's to make up for that fact."*

Can You Get an Edge in Obscure Sports?

- Betting discussions often suggest that the best opportunities are in **obscure sports or smaller leagues**.
- The logic is straightforward:
  - Bookmakers have less data
  - Their models may be less sophisticated
  - A bettor may be able to build a better model
- This argument is partly correct.
- But bookmakers largely anticipate this problem.

Bookmaker response

- **Higher margins** in obscure markets
- Often accompanied by **lower betting limits**

Implication

- It may be easier to find a predictive edge in smaller markets.
- But bookmakers have already priced this risk in.

Margins Across Time

- Betting discussions often suggest that the best strategy is to **bet early**, when odds may be misaligned with fundamentals.
- But the odds formation process usually works in the opposite direction.

Bookmaker behaviour

- When markets first open, bookmakers tend to be **more cautious**.
- They often start with **higher margins** to protect themselves.
- Early bettors may also be perceived as particularly confident, and therefore willing to accept worse odds.

Evidence

- Data from UK bookmakers for four major European football leagues (2022/23â2024/25, TheOdds-API.com).
- Average implied margins fall by about **0.4 percentage points** from two days before a match to the hour before kick-off.

Implication

- If you are not consistently identifying mispriced early odds, betting early likely means **paying a higher margin**.

Average margins fall before kick-off

Implied margins with UK bookmakers on four top European leagues from 2022/23 to 2024/25

Time to kick-off	*EPL*	*La Liga*	*Serie A*	*Bun-desliga*
48 hours	5.9%	6.5%	6.6%	6.4%
24 hours	5.8%	6.3%	6.4%	6.3%
18 hours	5.7%	6.3%	6.3%	6.2%
12 hours	5.7%	6.2%	6.3%	6.3%
6 hours	5.6%	6.1%	6.2%	6.1%
1 hour	5.5%	6.1%	6.2%	6.1%
10 minutes	5.5%	6.1%	6.2%	6.0%
1 minute	5.5%	6.1%	6.2%	6.0%

Supplementary Material in the Draft Book

- Chapter 19: Bigger Fields Mean Bigger Margins
- Chapter 20: Margins Across Markets and Time