The Price of Identity:

Overoptimism and Congruence Concerns

Kwabena Donkor[†] Lorenz Goette[‡] Maximilian W. Müller[§]

Eugen Dimant[¶] Michael Kurschilgen[‖]

April 16, 2026

Abstract

We study how social identity distorts both beliefs and preferences in investment decision-making. We ran framed field experiments in online sports betting with Kenyan and UK soccer fans during the 2021/22 Premier League season, eliciting subjective probabilities and incentivized bets on matches that sometimes involve a participant's supported team and sometimes do not. Supporters overestimate their team's win probability by 10–18 percent relative to neutral bettors and allocate about 20 percent more of their budget to bets on their own team than to otherwise similar matches where they are neutral. Embedding these data in a portfolio-choice model, we separate the role of distorted beliefs from that of identity-driven preferences. Distorted beliefs explain 30–44 percent of the excess investment in favored teams, while the remaining 56–70 percent reflects an aversion to profiting when one's team loses, equivalent to valuing payoffs from such outcomes by 17–27 percent less. Exploiting variation in both recent and season-long performance, we show that this preference to avoid "disloyal" gains is stronger when teams perform poorly. Using the estimated model, we calculate welfare losses from identity-driven distortions of roughly 0.5–2.7 percent of betting budgets, with the exact magnitude depending on whether identity preferences are seen as welfare-relevant or as mistakes.

Keywords: Social Identity, Belief Distortions, Investment, Field Experiment, Sports Betting

JEL Codes: D81, D91, G41, C93, Z10

We thank Yan Chen, Stefano DellaVigna, Matthew Gentzkow, Andreas Grunewald, Florian Hett, Alex Imas, Muriel Niederle, Ted O'Donoghue, Suanna Oh, Yesim Orhun, Moses Shayo, Ferdinand von Siemens, and seminar participants at Stanford GSB, ERINN, SITE, Tsinghua BEAT 2023, ASSA 2023, briq, Cornell, Frankfurt, Fribourg, Lausanne, MIT Sloan, Northwestern Kellogg, NUS, QME, UMich, Uppsala, and Wharton for valuable comments and discussions. Zihua Chen and Yiwei Fan provided excellent research assistance. The study's pre-registration is at aspredicted.org #73083. This study has received an exempt review by the National University of Singapore (ECSDERC-2021-12). Financial support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2126/1- 390838866 and by the University of Pennsylvania is gratefully acknowledged. Maximilian W. Müller acknowledges funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

[†]Stanford GSB (Email: donkor@stanford.edu).

[‡]National University of Singapore (Email: ecslfg@nus.edu.sg).

[§]Toulouse School of Economics (Email: maximilian.muller@tse-fr.eu).

[¶]University of Pennsylvania (Email: edimant@sas.upenn.edu).

[‖]UniDistance Suisse (Email: michael.kurschilgen@unidistance.ch).

1 Introduction

When group identity distorts economic decisions, how much of the distortion reflects what people believe and how much reflects what they prefer? The answer has direct policy implications.[1] If identity distorts decisions primarily through biased beliefs, correcting misperceptions could be sufficient to reduce the distortion. If identity also reshapes preferences, information alone will not suffice. Evidence that both channels may operate is widespread: investors overweight domestic equities despite diversification benefits (Foad, 2012; Huberman, 2015), consumers buy goods (food categories) that signal group membership (Atkin et al., 2021), and partisans evaluate identical economic data differently depending on which party holds office (Meeuwis et al., 2022). Yet, a common feature of such studies is that in each setting, beliefs and preferences are entangled. No study has jointly measured and separated the two channels in a single decision environment. Empirically separating beliefs from preferences is difficult in typical financial and consumer markets. Identities are diffuse, beliefs are hard to measure, and returns rarely vary exogenously.

A clean decomposition requires three features: identity affiliations that are measurable, stable over time, and exogenous to the decision context; direct elicitation of beliefs about outcomes with known ex-post realizations; and exogenous variation in returns to identify preferences separately from beliefs. Our paper provides the first joint decomposition of the belief and preference channels of identity in an investment context. We develop a portfolio-allocation model in which identity enters through two channels: distorted subjective beliefs about asset returns and an *identity-incongruence tax*, a shadow tax that captures how individuals undervalue payoffs from identity-incongruent sources (Akerlof and Kranton, 2000; Benjamin et al., 2010; Chen and Chen, 2011; Akerlof, 2016; Shayo, 2020; Henkel and Zimpelmann, 2023; Bénabou and Henkel, 2025). The model predicts that individuals overinvest in identity-congruent assets even after accounting for beliefs, risk preferences, and expected returns.

We test these predictions in framed field experiments (Harrison and List, 2004) with 2,410 Premier League soccer fans (802 from Kenya and 1,608 from the UK) who collectively placed nearly 40,000 bets during the 2021/22 season.[2][3] For each match, participants reported subjective probabilities for each outcome and then allocated a fixed budget across those outcomes at randomly assigned odds. The design keeps the investment problem constant: participants make the same type of financial decision whether or not identity is at stake. The randomized odds generate within-person and between-person variation in returns independent of identity, beliefs, or risk preferences.

Sports betting provides a setting with all three features needed for a clean decomposition: measurable identity affiliations, observable outcomes with probability benchmarks, and exogenous variation in returns. Team allegiances are long-standing, salient, and exogenous to match-level decisions. Match outcomes are publicly observable and have well-defined probabilities benchmarked by bookmaker odds.

[1] Consider, for example, discrimination: understanding whether behavior reflects taste-based motives or inaccurate beliefs has distinct policy and welfare implications (Bohren et al., 2023).

[2] We use bookmakers' predictions as forecast benchmarks for match outcomes. We compare participants' subjective beliefs against these benchmarks and assess distortions in beliefs related to fan identity. Having this benchmark is a major advantage of the experimental context, rarely available in settings such as reactions to global macroeconomic shocks.

[3] This paper substantively expands upon Morewedge et al. (2018), who also studied hedging in sports settings and found that people are reluctant to hedge when outcomes affect their identity. Unlike Morewedge et al. (2018), we directly measure belief distortions and use a structural framework to decompose the separate contributions of identity-specific beliefs and identity-driven preferences to identity-conforming behavior.

Betting markets post explicit returns that participants understand and respond to, and the experimental design randomly varies these returns across participants and matches, generating exogenous variation in the price of identity-congruent investment. Importantly, sports betting is not merely a laboratory curiosity. Legalized sports betting in the U.S. crowds out financial investments nearly dollar-for-dollar, with little effect on lottery or other gambling expenditures (Baker et al., 2024), and individuals in developing countries use sports betting strategically to finance lumpy expenditures (Herskowitz, 2021).[4] The mechanisms documented here (distorted beliefs and identity-driven preferences) are therefore likely to operate in many financial decisions involving identity-linked options, from domestic versus foreign assets to politically aligned versus neutral brands.

The reduced-form results reveal large identity-driven distortions. Supporters allocate about 20 percent more of their budget to bets on their team winning than to otherwise equivalent neutral matches, a difference we term the *affinity–neutrality investment gap*. Supporters also assess their team's win probability as 10–18 percent higher than neutral bettors do. These patterns persist after controlling for individual and match fixed effects. A semi-structural analysis that incorporates individual risk preferences rejects the hypothesis that, conditional on subjective beliefs, odds, and risk aversion, supporters and neutral bettors choose portfolios in the same way. Identity affects wagers over and above its effect on beliefs.

The structural estimation yields identity-incongruence tax rates of 17–27 percent: supporters discount the utility from payoffs associated with their team losing by this amount. Decomposing the affinity–neutrality investment gap, belief distortions account for 30–44 percent and identity-driven preferences account for the remaining 56–70 percent. The identity-incongruence tax is larger for supporters of underperforming teams. Both short-term losses and lower season-long standings intensify identity concerns, consistent with the *identity-threat response* theorized by Tajfel and Turner (1979). Using the estimated model, we calculate welfare losses from identity-driven distortions of 0.5–2.7 percent of betting budgets, with the magnitude depending on whether identity preferences are treated as welfare-relevant or as mistakes (Chetty, 2009; Bernheim and Taubinsky, 2019).

Our main contribution is to develop and implement a framework that separately estimates the effects of identity on beliefs and preferences in an investment context, providing a disciplined measure of the strength of each channel. A large literature documents that group membership shapes beliefs and expectations (Tajfel, 1970; Hogg, 1992; Chen and Li, 2009; Goette et al., 2006, 2012; Dimant, 2023). Partisan identity affects how people form beliefs about economic conditions (Bauer et al., 2023) and distorts perceptions of the environment (Cacault and Grieder, 2019; Bicchieri et al., 2023). Identity also shapes views on climate change, vaccines, and social norms (Egan and Mullin, 2017; Kurschilgen and Marcin, 2019; Allcott et al., 2020; Depetris-Chauvin et al., 2020; Cassan et al., 2021; Gadarian et al., 2021; Kurschilgen, 2023; Dimant, 2023; Dimant et al., 2024; Nunn, 2022; Gelfand et al., 2024). In sports, Radzevick and Moore (2008), Massey et al. (2011), and Simmons and Massey (2012) show that overoptimism is not cheap talk and persists despite accuracy incentives and learning opportunities. A smaller body of work demonstrates that identity preferences, holding beliefs fixed or assuming rational expectations, affect decisions such as sports betting, investment, occupational

[4]Because of the relatively controlled environment of sports, this setting is frequently used to study decision-making phenomena such as the hot hand fallacy, corruption, identity, and investment decisions (Wolfers, 2006; Miller and Sanjurjo, 2018; Depetris-Chauvin et al., 2020; Baker et al., 2024). With a global market size of USD 235 billion in 2022 (Statista, 2024) and soccer accounting for about a quarter of the market (Grand View Research, 2023), sports betting is itself an economically significant context.

choice, and food consumption (Paul and Weinbach, 2009; Morewedge et al., 2018; Atkin et al., 2021; Oh, 2023). Prior research has thus examined either belief distortions or identity-driven preferences, but not both jointly within a single framework. By combining field experiments that elicit both beliefs and investment decisions with a structural model of portfolio allocation, we can first use a semi-structural test to assess whether identity-related preferences exist after accounting for heterogeneity in beliefs and risk aversion, and then quantify the relative importance of each channel. This decomposition contributes to a growing literature on structural estimation of social and identity concerns in economic decision-making (DellaVigna et al., 2012, 2016; Karing, 2018; Donkor, 2021; Butera et al., 2022). Using a common metric to measure identity preferences allows for the comparison of identity concerns across groups, settings, and time. This methodological contribution thus can be used to make progress on questions of relevance to the understanding of identity and we use our methodology to illustrate the promise of our approach in two ways.

First, we use it for an empirical assessment of the identity-threat response (Tajfel and Turner, 1979): the hypothesis that individuals reinforce their commitment to a social identity when they perceive it as under threat. We exploit variation in short-term success (whether a team won its previous match) and longer-term success (final league standings from the prior season) to study how performance shapes identity concerns. Underperformance correlates with a systematically higher identity-incongruence tax. This pattern is consistent with individuals responding to identity threats by reinforcing in-group attachments. Prior tests of the identity-threat response have focused primarily on out-group hostility and intergroup conflict (Tajfel et al., 1971; Goette et al., 2006; Charness and Chen, 2020; Dimant, 2023); the results here extend the mechanism to the valuation of private financial payoffs from identity-incongruent outcomes.

Second, we use it to conduct a welfare analysis, contributing to the behavioral public economics literature on measuring the costs of decision-making biases (Bernheim and Taubinsky, 2019; Allcott and Taubinsky, 2015). By separately removing belief distortions and identity preferences in counterfactual simulations, we provide the first welfare estimates of identity-driven distortions in an investment context. The welfare consequences depend critically on whether identity preferences are treated as part of true utility or as mistakes, a normative choice that determines both the magnitude of welfare losses and the appropriate policy response. A central insight from Allcott and Taubinsky (2015) is that heterogeneity in behavioral biases raises aggregate welfare costs because losses are convex in the magnitude of the bias. The estimates here provide a direct illustration: in the UK sample, moving from a uniform identity-incongruence tax to team-specific estimates increases the welfare loss from 0.4 to 1.9 percent of the betting budget, even though the supporter-weighted mean tax is nearly identical. This finding underscores that policies targeting identity-driven distortions cannot rely on average bias as a sufficient statistic for welfare. These estimates also speak to the rapid expansion of legalized sports betting in the U.S., where 38 states had legalized some form of sports betting by 2023, and Americans wagered $121 billion (American Gaming Association, T., 2024; Baker et al., 2024; Hollenbeck et al., 2024). Unlike lottery participants, who are not systematically overoptimistic about winnings (Lockwood et al., 2021), sports bettors exhibit identity-driven overoptimism and overinvest in supported teams. A back-of-the-envelope calculation suggests that identity-driven distortions generate aggregate welfare losses for bettors between 120–160 million USD annually in the U.S. alone. These patterns suggest that effective regulation of sports betting markets should account for both belief distortions and identity-driven preferences.

The rest of the paper proceeds as follows. Section 2 develops the portfolio-allocation model, intro-

duces the identity-incongruence tax, and derives the decomposition of the affinity–neutrality investment gap into belief and preference components. Section 3 describes the experimental design and the two samples. Section 4 documents the reduced-form evidence on identity-driven betting and belief distortions. Section 5 tests whether identity affects portfolio allocations after accounting for subjective beliefs, odds, and risk preferences. Section 6 estimates the identity-incongruence tax, decomposes the investment gap, examines heterogeneity across teams and performance, and calculates welfare losses. Section 7 discusses alternative interpretations and policy implications. Section 8 concludes.

2 Theoretical Framework

We develop a portfolio allocation model that incorporates individuals' affinities toward specific assets that influence their investment decisions.[5] These affinities may arise from various sources, such as preferential investment in an employer's stock, personal entrepreneurial ventures, patriotic support for domestic firms, or alignment with a company's social or environmental mission. Such affinity-related preferences systematically affect investment behavior, leading to choices that diverge from those based solely on financial returns.

Consider an individual i deciding on how to allocate a fixed budget B across m mutually exclusive investment options in set J.[6] Let $j = 1, 2, \ldots, m$ index the investment opportunities in J, analogous to distinct states in the Arrow-Debreu setting. Each investment option j corresponds to a return θ_j that is realized with probability $p_{i,j}$. Since the returns are mutually exclusive, the probabilities sum to one: $\sum_{j=1}^{m} p_{i,j} = 1$. If individual i allocates $x_{i,j}$ to asset j, her return would be $\theta_j x_{i,j}$ with probability $p_{i,j}$. Budget constraints require that $\sum_{j=1}^{m} x_{i,j} \leq B$. We assume that individuals maximize subjective expected utility, where utility from investing in asset j is $u(\theta_j x_{i,j})$, with $u' > 0$ and $u'' < 0$. For simplicity, we classify the set J into two categories: $j = 1$, representing a specific asset of interest, and $j = J_{-1}$, representing a composite asset of the remaining options. Their corresponding marginal returns and beliefs are $\{\theta_1, p_{i,1}\}$ and $\{\theta_{J_{-1}}, p_{i,J_{-1}}\}$, respectively.

For a neutral investor (denoted as N) who has no affinity for any asset, the decision problem is:

$$\text{Max}_{x_{i,1}, x_{i,J_{-1}}} U_i^N = p_{i,1} u(\theta_1 x_{i,1}) + p_{i,J_{-1}} u(\theta_{J_{-1}} x_{i,J_{-1}}). \quad (1)$$

Now, consider an "identity" investor (denoted as H) who has an affinity for a specific asset in the pool of investment options. Suppose asset $j = 1$ is relabeled as $j = h$, and individual i has an affinity for what h symbolizes. This affinity introduces two effects. First, the investor's subjective probability of receiving a return from asset h may be distorted to $\tilde{p}_{i,h}$, which may differ from the neutral probability $p_{i,1}$. Second, payoffs from identity-incongruent assets (assets in J_{-h}) are discounted by a factor of $1-\alpha$,

[5] Arrow-Debreu securities provide a foundational structure for understanding how assets can be priced and allocated in a complete market with mutually exclusive outcomes. The model remains a cornerstone of economic theory, offering a clear and structured way to think about uncertainty, risk, and asset pricing.

[6] Examples of mutually exclusive investment options include betting on sports match results, investing in weather derivative securities, and betting on political election outcomes.

where $0 \leq \alpha < 1$.[7] The decision problem for an identity investor is expressed as:

$$\text{Max}_{x_{i,h}, x_{i,J_{-h}}} U_i^H \quad = \quad \tilde{p}_{i,h} u(\theta_h x_{i,h}) + (1-\alpha)\tilde{p}_{i,J_{-h}} u(\theta_{J_{-h}} x_{i,J_{-h}}). \tag{2}$$

The identity-incongruence tax α captures the concept of "source dependence" (Loewenstein and Issacharoff, 1994), the principle that people value outcomes differently based on their source. Objects associated with negative sources have been shown to induce disgust and reduce perceived value. For example, Stellar and Willer (2014) show that money from immoral sources is seen as morally contaminating and less desirable. In the model, α is a reduced-form parameter that encapsulates source-dependence and all other motives behind how identity drives preferences, reflected as the reduced desirability of money from identity-incongruent sources.[8] We place no structure on the belief formation process in our model because of the lack of consensus in the literature on the underlying mechanisms.[9] However, empirical evidence suggests that individuals often exhibit overoptimism regarding the returns to identity-congruent assets, a well-documented phenomenon.[10]

The model predicts the following investor behavior:

Proposition 1. *If Investor i has an affinity towards a particular asset (say asset h), then, independent of potentially identity-distorted beliefs, the investor's optimal budget allocation toward asset h will be greater when identity preferences are considered* $(0 < \alpha < 1)$ *compared to when they are not* $(\alpha = 0)$. *Formally,* $x_{i,h}(0 < \alpha < 1) \geq x_{i,h}(\alpha = 0)$.

Proposition 1 asserts that holding beliefs, risk preferences, and returns constant, identity preferences induce investors to allocate more to identity-congruent assets than neutral assets. The proof is provided in the Appendix section 8. Notably, this result contrasts with the notion of emotional hedging, where individuals might underinvest in identity-congruent assets to mitigate undesirable outcomes, implying $\alpha < 0$. We discuss this alternative interpretation in section 7.2.

An Application: Affinity-Neutrality Investment Gap Under CARA Utility

To provide a concrete application of how investment behavior may differ between neutral and identity investors because of distortions in beliefs or preferences, we impose more structure on our model. We assume that investors have CARA utility. Thus, the utility from the returns for wagering or investing in potential outcome j is $u(\theta_j x_{i,j}) = -e^{-r_i \theta_j x_{i,j}}$, where $r_i > 0$ denotes individual i's risk aversion. The

[7] This can be thought of as a specific formulation of state-dependent preferences, where we assume that individuals expect the marginal utility of an extra dollar in an identity-incongruent state to be lower than in a neutral state.

[8] We could also allow for a utility boost η from the team winning, which would enter equation (2) with weight $\tilde{p}_{i,h}$. As the model captures distortions in the perceived utility from winning and losing separately, this would not alter any of the portfolio decisions.

[9] For example, according to Bénabou and Tirole (2016), belief distortion may be due to motivated beliefs in the form of overconfidence, wishful thinking, and willful blindness. Other types of belief distortions are groupthink, team morale, and market exuberance.

[10] A large strand of literature in psychology and economics has documented that individuals systematically overestimate the likelihood of events perceived as positive (and underestimate the likelihood of negative events). Overoptimism has also been documented across many domains, including health (Weinstein, 1980), work (Hoch, 1985), marriage (Baker and Emery, 1993), life expectancy (Puri and Robinson, 2007), and sports (Radzevick and Moore, 2008). Studying overoptimism in the context of American football, Massey et al. (2011) and Simmons and Massey (2012) provide evidence that stated overoptimism is not just cheap talk, and that it persists despite accuracy incentives and opportunities to learn from real-world feedback.

neutral investor's optimal budget allocation to asset h is:

$$x_{i,h}^{N} = \frac{1}{\theta_h + \theta_{J-h}} \left[B\theta_{J-h} + \frac{1}{r_i} \ln \left(\frac{p_{i,h}\theta_h}{p_{i,J-h}\theta_{J-h}} \right) \right]. \tag{3}$$

For an identity investor, affinity towards asset (h) may introduce a distortion in beliefs ($\tilde{p}_{i,j}$) and a penalty ($\alpha > 0$) on consumption utility from returns from identity-incongruent assets. The identity investor's budget share assigned to h is:

$$x_{i,h}^{H} = \frac{1}{\theta_h + \theta_{J-h}} \left[B\theta_{J-h} + \frac{1}{r_i} \ln \left(\frac{\tilde{p}_{i,h}\theta_h}{(1-\alpha)\tilde{p}_{i,J-h}\theta_{J-h}} \right) \right]. \tag{4}$$

The difference between equations (3) and (4), is the affinity-neutrality investment gap (ANIG):

$$\underbrace{x_{i,h}^{H} - x_{i,h}^{N}}_{\text{Affinity-neutrality investment gap}} = \frac{1}{r_i(\theta_h + \theta_{J-h})} \left[\underbrace{\ln \left(\frac{\tilde{p}_{i,h}}{\tilde{p}_{i,J-h}} \right) - \ln \left(\frac{p_{i,h}}{p_{i,J-h}} \right)}_{\text{Share due to distorted beliefs}} + \underbrace{\ln \left(\frac{1}{1-\alpha} \right)}_{\text{Share due to identity-preferences}} \right].$$

This investment gap is driven by two components: (1) the change in log odds between neutral and distorted beliefs, and (2) identity preferences above and beyond the change in beliefs—the identity-incongruence tax α. Therefore, both overoptimism ($\tilde{p}_{i,h} > p_{i,h}$) and identity preferences ($\alpha > 0$) increase the ANIG.

The next section describes a field experiment in soccer betting that provides the data needed to decompose the ANIG empirically. Premier League matches present three mutually exclusive outcomes (j: home win, draw, and away win), and soccer fans typically have strong team affiliations, making betting on league games a natural setting for the framework above.

We conduct a lab-in-field experiment with soccer fans from the UK and Kenya, where participants with team affiliations engage in a betting game on several Premier League matches. Details of the experimental procedures and data collection methods are in the next section.

3 Experiment

Soccer betting offers a controlled environment for studying identity in investment decisions. Prior work has shown that soccer fans are reluctant to bet against their favored team (Morewedge et al., 2018; Kossuth et al., 2020) and overconfident in their predictions of match outcomes (Chegere et al., 2022).[11] The experimental design described below exploits this setting by varying betting odds exogenously, keeping the decision environment constant while changing the financial incentive to deviate from identity-congruent choices.

[11] Betting behavior has also been used to study corruption and savings behavior (Wolfers, 2006; Deutscher et al., 2025; Herskowitz, 2021).

3.1 Experimental Design

We exogenously vary the relative returns for betting on the home team's loss versus win. For each match, we obtained a measure of benchmark probabilities,[12] and then calculated seven sets of home (h) and away (a) odds (θ_h^z and θ_a^z): adjusted by a factor $z \in \{-2.5, -1.5, -.5, 0, +.5, +1.5, +2.5\}$,[13] where

$$z = p_a^0 * \theta_a^z - p_h^0 * \theta_h^z. \tag{5}$$

Before each matchday, we applied the formula (described in detail in the online Appendix section C.3) for the respective wedges z to the benchmark odds that we obtained from the German betting company "Tipico".[14] Bettors were not presented with the benchmark odds directly. Instead, we randomized odds at the individual-match-level, such that, for each match, respondents are shown a set of odds that is computed using a random draw of z.[15] Thus, each set of adjusted odds can be drawn with an equal probability ($\frac{1}{7}$). Participants were not informed that odds were randomized or how they related to bookmaker pricing. The interface displayed each set of odds alongside real-time calculations of potential returns for any given token allocation, so participants could evaluate the financial consequences of their choices without needing to calculate these themselves. In Table A4, we show an example of randomized odds (and the underlying change in assessed likelihood) for different sets of benchmark odds.

Our design provides two layers of exogenous variation in relative returns of betting on a win versus a loss: within-person variation in returns for the same team across matches and between-person variation in returns within matches. Therefore, our experimental variation in odds introduces exogenous variation in the monetary returns to investments independent of investors' identity (support of a team), their level of risk aversion, or subjective beliefs. This variation together with observing investors in both neutral and identity-relevant investment scenarios, allows us to test for the existence of identity preferences and estimate both investors' risk aversion and the strength of the identity tax.

3.2 Recruitment

We recruited 2,410 soccer fans from Kenya and the UK before the start of the Premier League Season 2021/22. In Kenya, we recruited through Ajua, a marketing firm that surveys young Kenyans for

[12] Bookmakers set odds based on assessed probabilities for home wins, draws, and away wins. Denote these assessed probabilities as $p_h^{BM}, p_d^{BM}, p_a^{BM}$, and the respective odds as θ_h^{BM}, θ_d^{BM}, and θ_a^{BM}. Expected payments from betting $\$x_j$ on outcome j would thus be equal to $p_j \theta_j^{BM} x_j$. If odds were fair this would imply that bookmakers set the following odds for each outcome j: $\theta_j^{BM} = \frac{1}{p_j^{BM}}$. In this case, odds are a perfect signal of assessed likelihoods as $p_j^{BM} = \frac{1}{\theta_j^{BM}}$ and thus $\sum_j p_j^{BM} = \sum_j \frac{1}{\theta_j^{BM}} = 1$. In practice, odds are not fair, and thus $p_j^{BM} \theta_j^{BM} < 1$. This implies that $p_j^{BM} < \frac{1}{\theta_j^{BM}}$ and as a consequence: $\sum_j \frac{1}{\theta_j^{BM}} > 1$, which is related to the bookmakers' profit margin. If bookmakers factor in a constant profit margin μ over what they perceive as fair odds $1/p_j^{BM}$, the underlying probabilities can be recovered as $\hat{p}_j = \frac{1/\theta_j^{BM}}{\sum_k 1/\theta_k^{BM}} = \frac{p_j^{BM}/(1-\mu)}{\sum_k p_k^{BM}/(1-\mu)} = p_j^{BM}$. We use this approximation when we infer bookmaker beliefs from a given set of odds.

[13] Since $p_a^0 = \frac{1}{\theta_a^0}$ and $p_h^0 = \frac{1}{\theta_h^0}$, the wedge is 0 for the benchmark odds.

[14] We used odds from a German betting site to minimize any potential bias in betting odds caused by biases in market demand driven by overoptimism on widely favored teams. While Page (2009) and Flepp et al. (2016) found neither team allegiances nor fan sentiment to affect bookmakers' pricing, using odds from a different country where most bettors have no strong allegiances towards Premier League teams further minimizes the role of such allegiances.

[15] Our experimental variation can introduce arbitrage opportunities for participants. To exploit such opportunities, participants should bet on outcomes with inflated odds in our experiment. We do not find evidence for such arbitrage behavior, suggesting potential narrow-bracketing. We discuss this in more detail in section 7.2.

market research. UK participants were recruited via Prolific. In both cases, participants were screened by stating they were soccer fans and supported a Premier League Club. To get more details on their fan identity, we asked participants before the season about their favorite team. During the season, we invited our participants a few days before matchdays to place bets on three or four upcoming games.[16]

We minimized priming or experimenter demand concerns in several ways. We conducted the screening and pre-season surveys a month prior to the league starting. Surveys on betting were then fielded during the league season. During the screening survey, participants were not informed that they would be invited back to bet on matches in the upcoming season. As such, they were not aware that this pre-season survey would be part of a larger study and that supporting some Premier League team would be a requirement to participate in the study. In the "betting surveys" during the season, we neither referred to the pre-season surveys nor made any reference to knowing what team they supported. All teams involved in matches during the season were treated equally in the surveys.

3.3 Participants' Decision Tasks

For the surveys during the season, we incentivized participants by paying a base amount of \$0.5 in Kenya and \$0.6 in the UK for each matchday in which they participated. Additionally, for each matchday in which they participated, we offered a bonus based on participants' wagers, the odds, and the actual match outcomes. In all survey rounds, we gave each respondent 100 tokens to bet on each match (equivalent to £2.0 (or approximately \$2.5) in the UK and KSh100 (or \$1.0) in Kenya). Each matchday, respondents were asked to allocate 100 tokens per match across match outcomes for at least three different matches and up to five matches. At the end of each survey round, we randomly selected one match per participant to determine their bonus and immediately informed participants which match would affect their bonus. When placing bets, participants do not know which match would be payoff-relevant, but learned so after having made all betting decisions for each matchday.

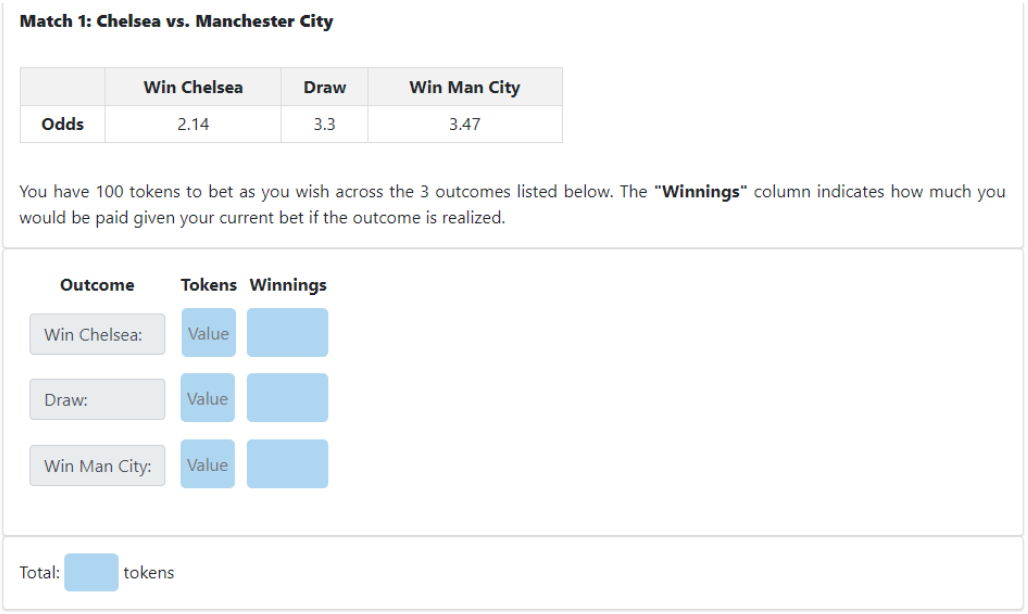
Under fair odds, participants can expect an average bonus of 100 tokens per matchday, amounting to KSh100 (\$1) in Kenya and £2 (approximately \$2.5) in the UK. Participants' earnings could exceed \$5 (£10) in Kenya (UK) if they bet 100 tokens on a high-odds outcome and it occurs.[17] We disbursed winnings a few days after each matchday.

Most participants completed the survey in about five minutes, which included two main questions about each match. First, we asked them to assign subjective likelihoods, ranging from 1% to 98% and totaling 100%, across the three-match outcomes. These likelihoods correspond to p_{ij} in our model. Second, we presented participants with odds for the three-match outcomes (θ_j in our model) and asked them to bet 100 tokens across these outcomes (x_h, x_d, x_a). Figure 1 shows how our interface displayed potential token winnings for each match outcome based on the given odds. For example, if a participant bet 40 tokens on "Win Chelsea", 40 on "Draw", and 20 on "Win Man City", it would show 86 tokens

[16] The pre-analysis plan specified Kenya and Singapore as study sites. Singapore was replaced with the UK because of difficulties recruiting a sufficient sample of Premier League supporters in Singapore. The final sample sizes also differ slightly from pre-specified targets due to variation in recruitment response rates across platforms.

[17] The maximum potential win was £75.4 in the UK and 3,770 KSh (\$37.7) in Kenya, which are large amounts, especially for an average duration of participation of 5 minutes. The average winnings also present meaningful incentives and are above what is typically paid for participation (Robinson, 2012; World Bank, 2022). Considering monthly basic minimum wages in Kenya as reported in Kenya National Bureau of Statistics (2022), 100 KSh are equivalent to 1.1% (.47%) of average monthly minimum wages in agriculture (in main urban areas) and thus 2.2 (.93) as much as the implied hourly wages (assuming standard weekly working hours of 45 hours). Recent research indicates that larger incentives are unlikely to further improve the quality of decisions in experiments (Enke et al., 2023).

Figure 1: Interface for Participants Placing Bets



Notes: This figure displays the interface for participants placing bets. They are first informed about the odds for a home win, a draw, and an away win. Below, participants can then type in how many tokens they would like to place on the 3 outcomes. Next to their allocated number of tokens under "Winnings", they will see how many tokens they would win in the case the respective outcome happens.

under "Winnings" for "Win Chelsea", 132 for "Draw", and 69 for "Win Man City". This simplified the task for participants by eliminating the need to calculate potential earnings.

Belief elicitation is not incentivized. The primary reason is to avoid creating a hedging opportunity: if belief reports were rewarded for accuracy, participants could strategically adjust reported probabilities to hedge against their betting positions. Neutral bettors' forecasts align closely with bookmaker odds (Section 4), suggesting that reports reflect genuine expectations rather than noise. Prior work also shows that identity-driven overoptimism in sports persists even under accuracy incentives (Massey et al., 2011), indicating that the biases documented here are unlikely to be artifacts of the elicitation method. To address residual measurement error, bookmaker odds serve as an instrument in the empirical strategy.

3.4 Sample Characteristics

Summary statistics about the participants are available in Table 1. Participants in Kenya are mostly male students in their twenties, whereas the UK sample is older and more balanced with 40% women. The UK sample is larger (N=1,608) than the Kenyan one (N=802), but participants in Kenya on average participated in twice as many matchdays (6.24 vs. 3.15), thus also betting on almost twice as many matches (24.31 vs. 12.46). Most respondents participated in multiple matchdays, 75% of respondents in each country participated in two or more matchdays.

Support for teams in our sample in Kenya (UK) is mostly concentrated among the following six teams: Manchester United with 36% (24%) of participants, Chelsea 31% (9%), Arsenal 18% (12%), Manchester City 9% (5%), Liverpool 4% (21%), and Tottenham 1% (8%). These six teams cover 99% of supporters in Kenya and 79% of supporters in the UK. An exact breakdown of supporters by teams is shown in the online Appendix table A1.

Table 1: Summary Statistics

	UK	Kenya
Number of Match Days (Survey Rounds)	6	11
Total Participants	1,608	802
Individual-Match Observations	20,034	19,499
Age	37.78	24.03
	(12.94)	(2.66)
Female (%)	40	12
Number of Match Days participated in (mean)	3.15	6.24
Number of bets (mean)	12.46	24.31
Supporter Matches (%)	12	28
Neutral Matches (%)	88	72
Number of years following Premier League (mean)	18.51	8.04
Number of years supporting favorite club (mean)	23.64	9.01

Notes: This table summarizes information about the survey respondents in the UK and Kenya, respectively. Standard deviations are reported in parentheses where relevant.

Team support is stable over time, as most participants have been supporting their favorite team for at least 10 years (in the Kenyan sample) or for 20 years or longer in the UK.[18] While the reasons participants have come to support a team are likely to be diverse (e.g., parents and relatives, peers, teams' success periods, jersey colors, or certain players), once they have picked a team, they most typically stick with it. This pattern is reflected in the ages at which the median participant started supporting their favorite team (15 years old in Kenya and 11 years in the UK). This observation is also similar for all teams, especially in the UK. The median years of support in Kenya (the UK) for the top 6 teams are as follows: 10 (20) years for Arsenal, 8.5 (18) years for Chelsea, 6 (20) years for Liverpool, 6 (15) years for Manchester City, 10 (20) years for Manchester United, and 6 (25) years for Tottenham Hotspur.

4 Reduced-Form Analysis

Here we document how identity influences both portfolio allocations and beliefs about match outcomes.

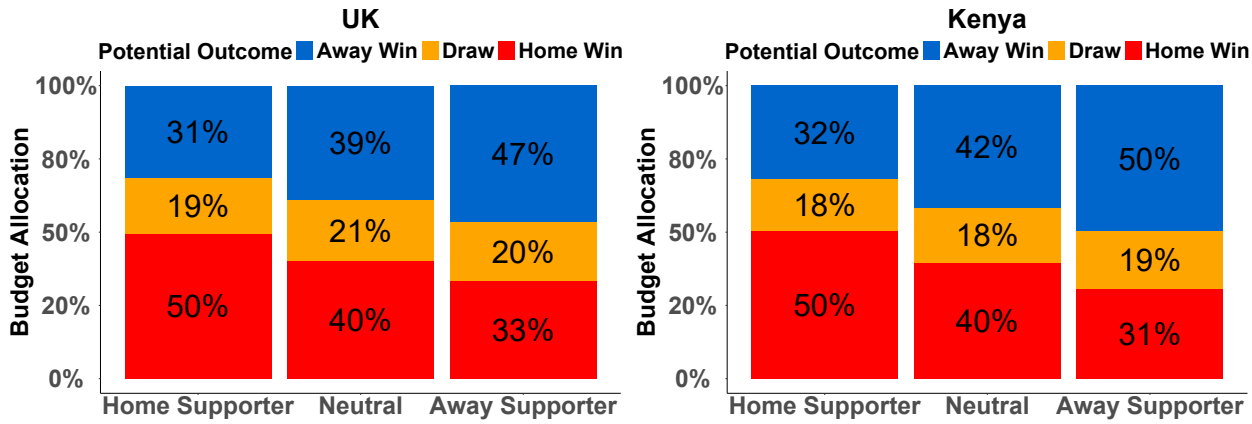
4.1 Portfolio Budget Allocation: Identity-Specific Betting

To assess the impact of identity on investment decisions, we analyze betting patterns among participants in the UK and Kenya. Figure 2 illustrates strikingly similar trends across both samples: home-team supporters allocate 10 percentage points more of their betting budget to a home win, while away-team supporters allocate eight percentage points less, compared to neutral bettors.

To isolate the effect of identity from individual and match-specific factors, we incorporate individual and match fixed effects in our regression analysis. Table 2 presents the adjusted results. Relative to neutral bettors, supporters of the home team allocate 8.5 percentage points more of their budget to a home win, while supporters of the away team reduce their bets on a home win by six percentage points.

[18] These differences across Kenya and the UK are mostly explained by differences in the average age in the two samples, as Kenyan participants are on average 24 and UK participants almost 38 years old.

Figure 2: Portfolio Allocations



Notes: The left and right panels above summarize how survey participants in the UK and Kenya, respectively, allocate their betting budget across the three potential outcomes of a soccer match (home-win, draw, away-win). The budget allocation is categorized by whether participants support the home or away team or are neutral across the teams in a match. The average betting shares are first calculated for each match and then averaged across all matches. We include graphs of the unconditional betting shares on home wins, draws, and away wins by home supporters, neutrals, and away supporters in the UK and in Kenya in Figures A1 and A2.

This pattern holds across both the UK and Kenyan samples. Similarly, supporters of the away team bet roughly eight percentage points more on an away win, while home-team supporters bet about six percentage points less.

Table 2: Portfolio Allocation

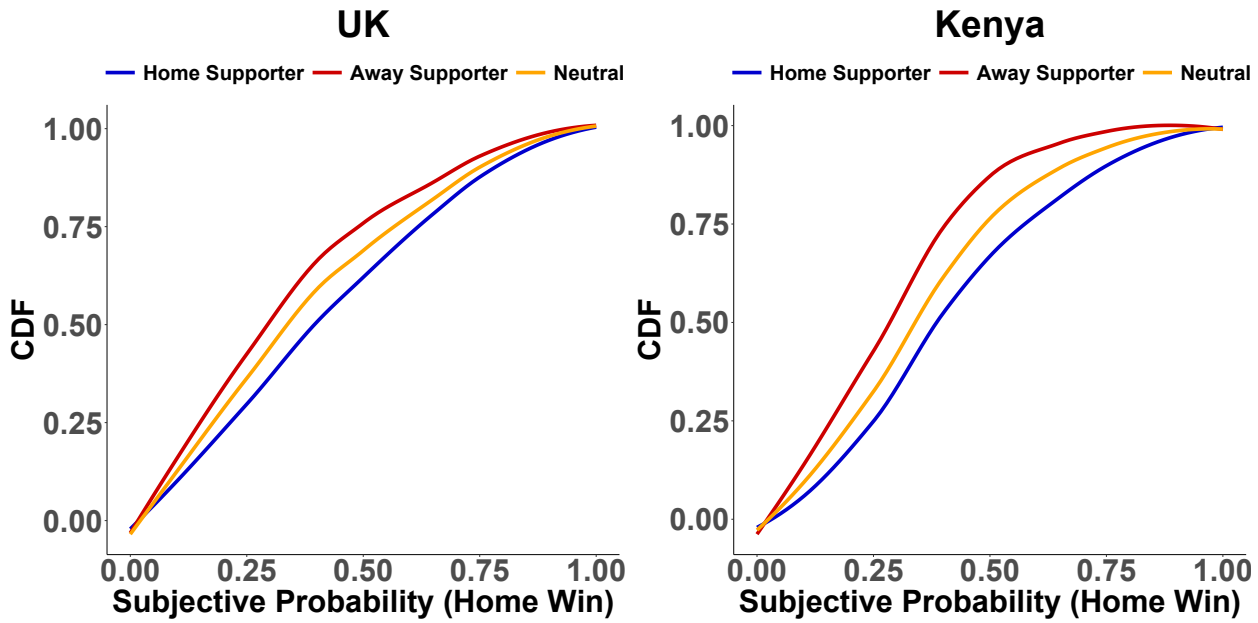
	Dependent Variable: Amount Bet on Outcome x_j			
	UK Home Win	UK Away Win	Kenya Home Win	Kenya Away Win
Home-Team Supporter	8.614^{***}	-6.032^{***}	8.456^{***}	-6.230^{***}
	(0.797)	(0.681)	(0.656)	(0.537)
Away-Team Supporter	-6.299^{***}	7.752^{***}	-5.993^{***}	7.638^{***}
	(0.606)	(0.702)	(0.540)	(0.629)
Outcome Mean				
Neutral Fan	38.71	40.33	39.4	41.82
Individual FE	Yes	Yes	Yes	Yes
Match FE	Yes	Yes	Yes	Yes
Observations	19,639	19,639	18,408	18,408
R^2	0.435	0.413	0.468	0.457

Notes: This table reports the average differences in survey participants' wagers on home-win and away-win outcomes for home-team and away-team supporters in the UK (columns (1) and (2)) and in Kenya (columns (3) and (4)). The reference group in each column is neutral fans. We account for both individual and match fixed effects in all columns. We exclude individuals for whom we estimate zero or negative levels of risk aversion (see section 5.2). We cluster our standard errors at the individual level and report the robust standard errors for our estimates in parentheses. The significance levels of our estimates are depicted as $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

4.2 Identity-Distorted Beliefs

We now examine how team allegiance distorts beliefs about match outcomes, testing for systematic overoptimism. Figure 3 presents the cumulative distribution functions (CDFs) of participants' subjec-

Figure 3: Subjective Beliefs (CDFs): Home-Team Win



Notes: The left and right panels above show the CDFs of survey participants' self-reported beliefs about "home team winning" of Premier League soccer matches for the UK and Kenya, respectively. CDFs are first calculated for each match and then averaged across all matches. The CDFs in each panel are categorized by whether the participants support the home (blue) or the away team (red) or are neutral (yellow) across the teams playing.

tive probabilities of a home-team win, comparing supporters and neutral bettors. Across both the UK and Kenya, supporters consistently overestimate their team's likelihood of winning relative to neutrals. This bias persists even after accounting for individual and match fixed effects (Table 3).

To assess the magnitude of this overoptimism, we compare neutral bettors' beliefs to bookmaker odds. In Kenya, neutral participants' predictions align closely with bookmakers, both estimating a 39% probability of a home-team win. In the UK, however, neutral bettors are slightly more pessimistic, assigning a 39% probability compared to bookmakers' 41%. Similarly, neutral bettors in both countries estimate the probability of an away-team victory at approximately 39% to 40%, which exceeds the bookmaker forecast of 36%.

Supporters, by contrast, exhibit a systematic bias in favor of their team. In the UK, home-team supporters predict their team's win probability to be 4.4 percentage points higher than neutral bettors do, while away-team supporters inflate their team's chances by 4.0 points. This effect is even stronger in Kenya, where home and away supporters exaggerate their team's winning probability by 7.0 and 5.2 percentage points, respectively. Across both countries, supporters' forecasts exceed neutral bettors' projections by 10% to 18%. Beyond inflating their team's chances of winning, supporters also systematically underestimate the probability of their team losing, both at home and away. This tendency is particularly pronounced for away-team supporters, who deviate most from bookmaker odds. These findings provide strong evidence of systematic overoptimism, where identity-driven biases shape individuals' beliefs about match outcomes, leading them to perceive their team as more likely to win than neutral bettors or bookmakers' benchmarks suggest.

Table 3: Subjective Beliefs

	Dependent Variable: Probability of Outcome j			
	UK Home Win	UK Away Win	Kenya Home Win	Kenya Away Win
Home-Team Supporter	0.044^{***}	-0.033^{***}	0.070^{***}	-0.049^{***}
	(0.005)	(0.004)	(0.005)	(0.004)
Away-Team Supporter	-0.032^{***}	0.040^{***}	-0.041^{***}	0.052^{***}
	(0.004)	(0.005)	(0.004)	(0.005)
Outcome Mean				
Neutral Fan	0.39	0.39	0.39	0.4
Bookmaker	0.41	0.36	0.39	0.36
Individual FE	Yes	Yes	Yes	Yes
Match FE	Yes	Yes	Yes	Yes
Observations	19,639	19,639	18,408	18,408
R^2	0.695	0.653	0.527	0.503

Notes: This table reports average differences in survey participants' reported beliefs about home-win and away-win outcomes for home-team and away-team supporters in the UK (columns (1) and (2)) and in Kenya (columns (3) and (4)). The reference group in each column is neutral fans. We account for both individual and match fixed effects in all columns. We exclude individuals for whom we estimate zero or negative levels of risk aversion (see section 5.2). We cluster our standard errors at the individual level and report the robust standard errors for our estimates in parentheses. The significance levels of our estimates are depicted as $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

5 Semi-Structural Analysis: Differences in Portfolio Allocation

Standard models of rational investment predict that portfolio decisions depend solely on subjective expectations of asset returns and individual risk preferences. If investors behave rationally, then after accounting for these factors, we should observe no systematic differences in portfolio allocation. This section empirically tests whether fan identity influences betting behavior beyond what can be explained by beliefs, betting odds, and risk preferences alone. We test the null hypothesis: conditional on subjective beliefs, odds, and individual risk preferences, team allegiance does not influence investment decisions.

We rely on the theoretical model outlined in Section 2, where we assume that individuals exhibit constant absolute risk aversion (CARA) utility to test the null hypothesis.[19] In the model, individuals maximize their subjective expected utility when placing bets. Given a match between two teams, individual i must decide how to allocate a fixed budget across three possible outcomes: h (home win), a (away win), or d (draw). If identity does not influence decision-making beyond subjective beliefs and risk preferences, then bettors with identical expectations and risk tolerance should allocate their budgets similarly, regardless of whether they support a team or remain neutral.

5.1 Portfolio Allocation Without Identity Concerns

If the null hypothesis holds, that is, if team allegiance does not affect betting behavior beyond subjective beliefs and risk preferences, then a neutral bettor's optimal budget allocation follows directly from the first-order conditions (FOCs) derived from our CARA utility model. Specifically, the optimal budget

[19] The choice of CARA utility is without loss of generality. The online Appendix section A presents the structural equations under CRRA utility. We select CARA utility because it yields first-order conditions that are easier to interpret and directly visualize using our experimental data, as shown in Figures 4 and A5.

shares for betting on a home win and an away win are given by:

$$x_{i,h}^{N} = \frac{\theta_d\theta_a}{\Theta} + \frac{1}{r_i}\underbrace{\left[\frac{(\theta_d+\theta_a)}{\Theta}\text{Ln}\left(\frac{p_{i,h}\theta_h}{p_{i,d}\theta_d}\right) - \frac{\theta_d}{\Theta}\text{Ln}\left(\frac{p_{i,a}\theta_a}{p_{i,d}\theta_d}\right)\right]}_{\Delta P_{i,h}} \tag{6}$$

$$x_{i,a}^{N} = \frac{\theta_d\theta_h}{\Theta} + \frac{1}{r_i}\underbrace{\left[\frac{(\theta_d+\theta_h)}{\Theta}\text{Ln}\left(\frac{p_{i,a}\theta_a}{p_{i,d}\theta_d}\right) - \frac{\theta_d}{\Theta}\text{Ln}\left(\frac{p_{i,h}\theta_h}{p_{i,d}\theta_d}\right)\right]}_{\Delta P_{i,a}}. \tag{7}$$

In these equations, Θ is defined as: $\Theta = \theta_a\theta_h + \theta_d\theta_h + \theta_d\theta_a$, and we normalize the total investment budget to one.[20] The first-order conditions provide an intuitive mechanism for budget allocation. The terms $\frac{\theta_d\theta_a}{\Theta}$ in equation (6) and $\frac{\theta_d\theta_h}{\Theta}$ in equation (7) represent the baseline allocation derived solely from the betting odds. They capture the "insurance" effect: investors naturally spread their budget across outcomes to hedge against risk.

Each equation includes a second term scaled by the inverse of the individual's CARA risk parameter r_i. This term reflects the sensitivity of the allocation to the difference in expected marginal returns between outcomes. For example, in equation (6), the expression $\Delta P_{i,h}$ captures the odds-adjusted difference in the logarithm of the subjective expected marginal returns between a home win and an away win, using draw as a benchmark. In essence, if an individual expects that a home win will offer a substantially higher marginal return relative to an away win, then $\Delta P_{i,h}$ will be larger, prompting a higher budget allocation to the home win. Similarly, equation (7) defines $\Delta P_{i,a}$ for away win.

The scaling by $1/r_i$ of the ΔPs indicates that the less risk-averse an individual is (i.e., the smaller r_i is), the more responsive their allocation is to these differences in expected returns. The ratios inside the logarithms combine both the subjective probabilities $p_{i,j}$ and the exogenously provided odds θ_j. These terms adjust for the fact that the attractiveness of a bet depends on both how likely an individual believes an outcome is and the payout offered by the odds.

Although an increase in the odds for a home win (θ_h) directly boosts the potential return from that outcome, it also triggers a redistribution of the budget. This occurs because of the diminishing marginal utility of gains. A rational, risk-averse investor will shift funds toward outcomes with relatively higher marginal utility. In other words, the investor insures against risk by balancing the enhanced return with the utility loss from concentrating too much on one outcome.

We use these baseline FOCs (equations (6) and (7)) to establish the neutral portfolio allocation. By comparing these predictions to actual betting behavior, we can identify whether deviations arise due to identity-driven effects rather than differences in subjective beliefs or risk preferences. In our experiment, we elicit subjective beliefs directly and vary betting odds exogenously. Thus, the only parameter we need to estimate from observed behavior is the CARA risk preference parameter r_i.

[20] In these expressions we remain agnostic to the possibility that a bettor's subjective probabilities could be distorted by identity (denoted as $\tilde{p}_{i,j}$).

5.2 Estimating Individual Risk-Preference

We estimate each individual's CARA risk preference parameter, r_i, using individual-level panel data on betting decisions. Because equations (6) and (7) jointly determine r_i at the match level, variation in betting decisions and beliefs, within and across matches, identifies r_i. Using individual-level panel data, we estimate equations (6) and (7) simultaneously in a linear regression framework as follows:

$$x_{j,m} - \frac{\theta_{d,m}\theta_{-j,m}}{\Theta_m} = \beta \times \Delta P_{j,m} + \varepsilon_{j,m}. \tag{8}$$

$x_{j,m}$ denotes the budget share that individual i allocates to outcome j (either a home or away win) in match m. The term $\frac{\theta_{d,m}\theta_{-j,m}}{\Theta_m}$ represents the baseline allocation predicted by the odds, with $\theta_{-j,m}$ denoting the odds of an away (home) win when j represents a home (away) win, and $\Theta_m = \theta_{a,m}\theta_{h,m} + \theta_{d,m}\theta_{h,m} + \theta_{d,m}\theta_{a,m}$. $\Delta P_{j,m}$ captures the odds-adjusted difference in the logarithm of subjective expected marginal returns across outcomes. $\varepsilon_{j,m}$ is an i.i.d. residual error. The coefficient β is equal to $\frac{1}{r_i}$.

A potential concern is that respondents might misreport their subjective beliefs, which would bias our estimate of $\Delta P_{j,m}$ and, consequently, β (and r_i). We consider two types of measurement error. First, random errors in the reported subjective beliefs can attenuate the estimate of β, leading to an overestimation of risk preferences. We mitigate this problem by using bookmaker odds as an instrument, which helps correct for attenuation bias.[21]

Second, supporters might exaggerate their team's chances to create a favorable self-image, but still use more accurate beliefs when allocating their bets. The net bias on β depends on how such exaggerations affect the log differences in $\Delta P_{j,m}$. For instance, if a respondent inflates the probability of a home win ($p_{i,h}$), it increases the numerator in $\Delta P_{i,h}$ and biases the estimate of $\frac{1}{r_i}$ toward zero (resulting in an overestimate of r_i). Addressing this systematic error is more challenging, especially when fan identity may be a confounding factor. To overcome this challenge, we restrict our estimation to matches where individuals are neutral, thereby reducing the likelihood of systematic misreporting.

5.3 Empirical Test for Null Hypothesis

With all the structural components of the first-order conditions (FOCs) in hand, we now test whether fan identity systematically influences portfolio allocations. Our goal is to determine if team allegiance affects betting behavior beyond what can be explained by subjective beliefs, odds, and risk preferences. To do this, we estimate the following regression model:

$$\begin{aligned} x_{i,j,m} - \frac{\theta_{i,d,m}\theta_{i,-j,m}}{\Theta_{i,j,m}} - \hat{r}_i^{-1}\Delta\hat{P}_{i,j,m} = {} & \gamma_1 \text{Home-Team Supporter}_{i,m} + \\ & \gamma_2 \text{Away-Team Supporter}_{i,m} + \\ & I_i + \lambda_m + e_{i,j,m}. \end{aligned} \tag{9}$$

The outcome represents the deviation of the observed budget share $x_{i,j,m}$ (the budget share allocated

[21] In the IV first stage, we construct an alternative measure $\Delta P^{BM}_{j,m}$ using bookmaker beliefs. We then use $\Delta P^{BM}_{j,m}$ as an instrument for the potentially mismeasured $\Delta P_{i,j,m}$.

by individual i to outcome j in match m) from the baseline allocation predicted by our theoretical model for neutral bettors. The baseline allocation includes two parts: (1) The odds-driven component, $\frac{\theta_{i,d,m}\theta_{i,-j,m}}{\Theta_{i,j,m}}$, and (2) the adjustment for subjective beliefs and risk preferences, $\hat{r}_i^{-1}\Delta\hat{P}_{i,j,m}$, where $\hat{r}_i$ is the estimated CARA risk aversion parameter from our earlier regression (equation (8)).

The right-hand side contains indicator variables for whether individual i is a home-team supporter (Home-Team Supporter$_{i,m}$) or an away-team supporter (Away-Team Supporter$_{i,m}$) in match m. We include individual fixed effects (I_i) and match fixed effects (λ_m) to control for unobserved heterogeneity across individuals and match-specific factors that might influence betting behavior.

If the null hypothesis is correct, the adjustments for neutral behavior should fully explain the observed allocations, and the coefficients γ_1 and γ_2 will be statistically indistinguishable from zero. A statistically significant coefficient would imply that, beyond beliefs and risk preferences, identity directly impacts investment choices.

5.4 Semi-Structural Results

We find systematic differences in betting behavior based on fan identity after accounting for subjective beliefs, odds, and risk preferences. This challenges the hypothesis that all bettors, regardless of team allegiance, decide on portfolio allocations like neutral bettors.

Figure 4 visually illustrates the empirical components of the theoretical portfolio allocation for a home win (as given in equation (6)). The figure displays a binned scatter plot of home-win bets against the fitted values of $\Delta\hat{P}_{i,j}$, with separate markers for home-team supporters (blue), away-team supporters (red), and neutral bettors (yellow). Fitted lines by fan identity, whose slopes represent the inverse of the identity-specific CARA risk preference parameters, are also shown. Under the null hypothesis, these points and lines should overlap, indicating no systematic differences by team allegiance. Instead, the data reveal that home-team supporters consistently allocate a higher share of their budget to home wins compared to neutrals, while away-team supporters allocate lower shares. This pattern holds in both the UK and Kenyan samples.

The experimental design may allow for arbitrage opportunities. If participants were indifferent to identity and fully rational, they would exploit the inflated odds in the experiment by wagering their entire budget on the outcome with the highest odds, effectively appearing risk-neutral. However, we observe no evidence of such arbitrage behavior. This absence may stem from narrow-bracketing, a cognitive bias in which individuals focus on individual bets rather than the overall portfolio, which aligns with observed risk aversion for small stakes. An alternative explanation is that participants possess a noisy mental representation of the true investment opportunities (Khaw et al., 2021).

The positive slopes of the fitted lines by fan identity indicate that, on average, participants are risk-averse. Although the lines display distinct offsets (or wedges) by fan identity, their largely parallel orientation suggests that the level of risk aversion remains consistent across groups. Therefore, differences in risk preferences are unlikely to account for the systematic variation in betting behavior between neutral bettors, home-team supporters, and away-team supporters.

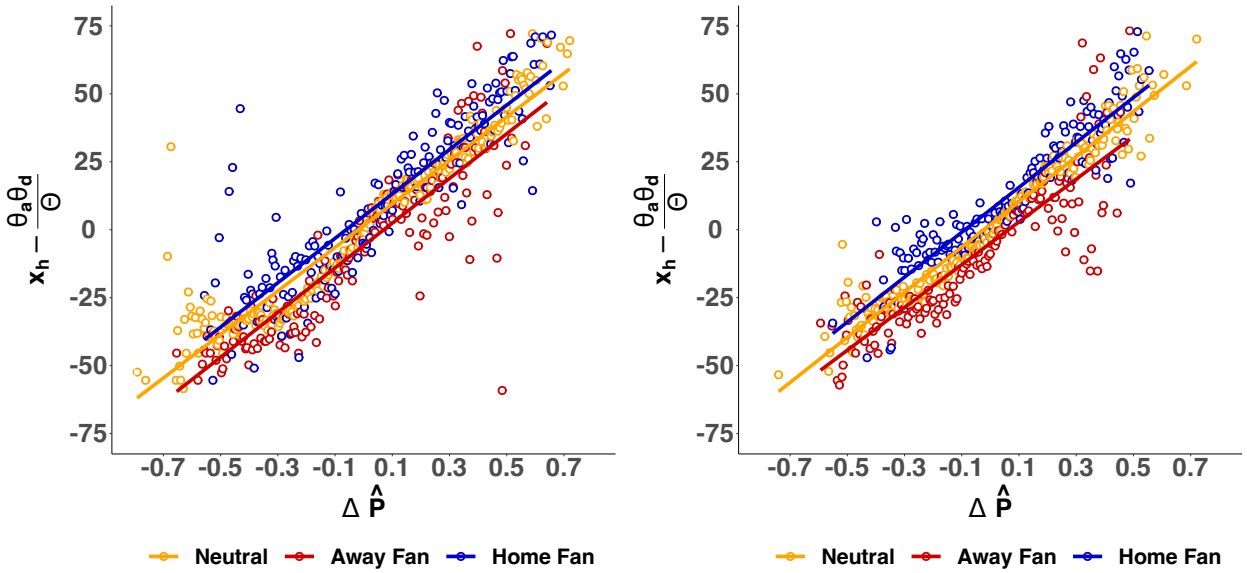
For a formal test of the null hypothesis, we estimate a two-stage least squares (TSLS) regression analogous to the graphical analysis in Figure 4, assuming homogeneous risk preferences. We instrument $\Delta P_{i,j,m}$ with $\Delta P^{BM}_{j,m}$ and include individual and match fixed effects. The results, presented in Table 4, show that the differences in portfolio allocations between home and away supporters versus neutral

Table 4: Semi-Structural Portfolio Allocation (Homogeneous Risk Preferences)

	Dependent Variable: $x - \frac{\theta_d \theta_j}{\Theta}$ (Bet minus ratio of odds)			
	UK Home Win	UK Away Win	Kenya Home Win	Kenya Away Win
Home-Team Supporter	3.482^{***}	-2.629^{***}	1.676^{***}	-1.953^{***}
	(0.744)	(0.664)	(0.592)	(0.532)
Away-Team Supporter	-2.762^{***}	3.470^{***}	-2.540^{***}	2.908^{***}
	(0.628)	(0.698)	(0.498)	(0.542)
$\Delta \hat{P}$	110.102^{***}	109.433^{***}	95.523^{***}	96.102^{***}
	(1.559)	(1.527)	(1.576)	(1.607)
CARA risk aversion (r)	0.009	0.009	0.01	0.01
Instrument for ΔP	ΔP^{BM}	ΔP^{BM}	ΔP^{BM}	ΔP^{BM}
Individual FE	Yes	Yes	Yes	Yes
Match FE	Yes	Yes	Yes	Yes
Observations	19,639	19,639	18,408	18,408
R^2	0.503	0.495	0.494	0.491

Notes: This table presents semi-structural estimates corresponding to the first-order conditions for neutral bets computed separately for both the UK (columns (1) and (2)) and for Kenya (columns (3) and (4)). In each column, we regress wagers on home-win (or away-win) x_j minus $\frac{\theta_d \theta_j}{\Theta}$ on indicators for home- and away-team supporters and ΔP. In each column, we assume homogeneous risk preferences across individuals with the CARA risk preference parameter estimated as the reciprocal of the coefficient on ΔP). We also instrument ΔP (computed using individual beliefs and the experimental odds) using a corresponding value of ΔP computed using bookmaker beliefs and the experimental odds (ΔP^{BM}). We exclude individuals for whom we estimate zero or negative levels of risk aversion (see section 5.2). We cluster our standard errors at the individual level and report the robust standard errors for our estimates in parentheses. The significance levels of our estimates are depicted as $*p < 0.10$, $**p < 0.05$, $***p < 0.01$.

Figure 4: Semi-structural Portfolio Allocation (Conditional on Beliefs)



Notes: The panels above show a scatter plot of actual wagers on home-win (minus the ratio of odds provided) against the model-computed wedge term ($\Delta\hat{P}$) of the first-order conditions of neutral bets on home-win for the UK and Kenya, respectively. The points in each panel are color-coded by supporter status: yellow for neutral bettors, blue for home-team supporters, and red for away-team supporters. The horizontal axis is $\Delta\hat{P}$, which captures the non-linear relationship of how beliefs and odds affect the amount wagered on a potential outcome.

bettors are statistically significant. The estimated homogeneous CARA risk preference parameter r is approximately 0.01 for both the UK and Kenyan samples, confirming that, on average, individuals are risk-averse.

In a next step, we allow for individual differences in risk aversion, r_i, as detailed in Section 5.2. Figure A3 shows the distribution of these estimates, with the Kenyan sample exhibiting an average r_i of 0.021 compared to 0.015 in the UK. When we incorporate these individual-specific risk preferences into our regression (equation (9)), the identity effects on portfolio allocation become even more pronounced. Table 5 reports these results, revealing that the differences in betting behavior based on fan identity roughly double once we account for heterogeneity in risk aversion. As a robustness check, we replace (impute) respondents' beliefs with bookmakers' beliefs and re-estimate the specifications in Table A3 in the Appendix. We find slightly larger differences in bets by fan identity. This finding suggests that the differences in bets are not driven by respondents reporting incorrect beliefs to the researcher while using more accurate beliefs when deciding on portfolio allocations.

6 Structural Analysis

We measure how supporters' and neutrals' portfolio choices when betting differ after accounting for subjective beliefs and risk preferences. The portfolio allocation model with identity concerns (equation (2)) suggests that individuals may undervalue (or underweight) the expected utility from gains from

Table 5: Semi-structural Portfolio Allocation (Heterogeneous Risk Preferences)

	Dependent Variable: $x - \frac{\theta_d \theta_j}{\Theta} - r_i^{-1}\Delta\hat{P}$			
	UK Home Win	UK Away Win	Kenya Home Win	Kenya Away Win
Home-Team Supporter	6.838^{***}	-5.000^{***}	6.662^{***}	-6.708^{***}
	(0.773)	(0.708)	(0.608)	(0.601)
Away-Team Supporter	-5.835^{***}	6.941^{***}	-6.605^{***}	6.925^{***}
	(0.671)	(0.708)	(0.589)	(0.604)
Instrument for ΔP	ΔP^{BM}	ΔP^{BM}	ΔP^{BM}	ΔP^{BM}
Individual FE	Yes	Yes	Yes	Yes
Match FE	Yes	Yes	Yes	Yes
Observations	19,639	19,639	18,408	18,408
R^2	0.162	0.166	0.101	0.106

Notes: This table presents semi-structural estimates corresponding to the first-order conditions for neutral bets computed separately for both the UK (columns (1), (2)) and for Kenya (columns (3), (4)). In each column we regress wagers on home-win (or away-win) x_j minus $\frac{\theta_d \theta_j}{\Theta} - r_i^{-1}\Delta\hat{P}$ on indicators for home- and away-team supporters. r_i^{-1} is the estimated individual-specific CARA risk preference parameter. $\Delta\hat{P}$ is the predicted value from an IV first-stage regression of $\Delta\hat{P}$ on ΔP^{BM}. We exclude individuals for whom we estimate zero or negative levels of risk aversion (see section 5.2). We cluster our standard errors at the individual level and report the robust standard errors for our estimates in parentheses. The significance levels of our estimates are depicted as $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

their supported team not winning a match.[22]

6.1 Portfolio Allocation with Identity Concerns

Risk-averse agents do not solely consider financial returns and risk when making investment decisions; they also factor in how well an outcome aligns with their identity (Proposition 1). In this context, fans care about betting on outcomes that are consistent with their team allegiance. For ease of exposition, we assume that individual i identifies with the home team. Under the CARA framework, the optimal budget allocations for an identity-driven bettor are given by:

$$x_{i,h}^H = \frac{\theta_d \theta_a}{\Theta} + \frac{1}{r_i}\Delta\tilde{P}_{i,h} + \underbrace{\mathrm{Ln}\,(1-\alpha)\, r_i^{-1} \times \frac{-(\theta_a+\theta_d)}{\Theta}}_{\text{Overinvestment}} \tag{10}$$

$$x_{i,a}^H = \frac{\theta_d \theta_h}{\Theta} + \frac{1}{r_i}\Delta\tilde{P}_{i,a} + \underbrace{\mathrm{Ln}\,(1-\alpha)\, r_i^{-1} \times \frac{\theta_d}{\Theta}}_{\text{Underinvestment}}. \tag{11}$$

[22] The specification of identity preferences as an incongruence tax α on utility leaves the curvature of the utility function unchanged. One might argue that identity changes the curvature of the utility function due to source preference (Chew and Sagi, 2008). Notice however, that changes in r_i due to identity would predict a sharply different pattern in Figure 4. In particular, changes in risk aversion would rotate the curves around $\Delta P = 0$ as predicted by equations (6) and (7). Our evidence lends no support to this prediction.

Table 6: Structural Estimates

	Homogeneous r		Heterogeneous r_i	
	UK	Kenya	UK	Kenya
α: Identity-incongruence tax rate	0.118^{***}	0.094^{***}	0.166^{***}	0.272^{***}
	(0.006)	(0.004)	(0.008)	(0.004)
r: CARA risk preference	0.009^{***}	0.010^{***}	-	-
	(0.000)	(0.000)	-	-
Individual-Outcome FE	Yes	Yes	Yes	Yes
Match-Outcome FE	Yes	Yes	Yes	Yes

Notes: This table reports estimates of the identity-incongruence tax rate that is homogeneous across supporter identity, that is, home-team or away-team supporters. The first two columns present estimates that assume homogeneous risk preferences, and the third and fourth columns account for individual-specific risk preferences. The estimating equation corresponds to the first-order conditions for supporter bets computed separately for the UK and Kenya. Individual and match-outcome fixed effects are accounted for in all columns. We report bootstrapped standard errors for our estimates in parentheses. The significance levels of our estimates are depicted as $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The term $\frac{\theta_d \theta_a}{\Theta}$ (or $\frac{\theta_d \theta_h}{\Theta}$) represents the baseline allocation derived from the betting odds, similar to neutral bettors. $\Delta\tilde{P}_{i,h}$ (and $\Delta\tilde{P}_{i,a}$) captures the odds-adjusted difference in the logarithm of subjective expected marginal returns. Note that these terms now reflect any distortions in beliefs due to identity. The additional term involving $\ln(1-\alpha)$ modifies the neutral allocation by penalizing the expected utility from outcomes that are incongruent with the fan's identity. For a home supporter, this adjustment leads to overinvestment in home wins (and, correspondingly, underinvestment in away wins).

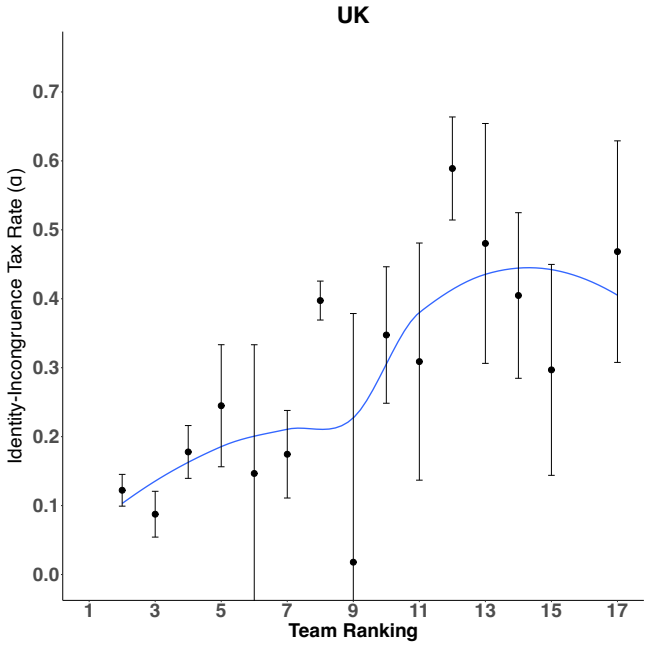
To quantify this identity tax, we estimate the parameter α using equations (10) and (11). We follow a similar estimation procedure as in the semi-structural analysis. The results, presented in Table 6, differentiate between the UK and Kenyan samples. Under the assumption of uniform risk preferences, we find an identity tax of approximately 11.8% for the UK and 9.4% for Kenya. In both samples, the homogeneous CARA risk parameter r is estimated to be around 0.01.

When we allow for individualized risk preferences (columns 3 and 4 of Table 6), the estimated identity tax increases markedly to 16.6% in the UK and 27.2% in Kenya. One interpretation for the higher tax estimates under heterogeneous risk preferences is that a significant share of individuals, particularly in Kenya, exhibit above-average risk aversion. For these individuals, deviating by a fixed amount from the neutral benchmark is particularly costly in utility terms, which is rationalized by a higher identity tax.

6.2 Heterogeneity in Identity Preferences

The framework allows us to examine how identity concerns vary across groups and over time. In particular, we test the "identity-threat response", the idea that individuals reinforce their social identity when they perceive it to be under threat (Tajfel and Turner, 1979). We explore this prediction by examining how team performance impacts identity preferences. In this setting, this theory predicts that poorer team performance should heighten identity concerns, resulting in a stronger identity-incongruence tax (α). In other words, when a team underperforms, fans should exhibit stronger

Figure 5: Heterogeneity in Identity-incongruence Tax Rate



Notes: This figure plots the estimates of the identity-incongruence tax rate for individual teams based on data from the 2021/22 season and with teams ranked by their final ranking at the end of the 2020/21 Premier League season. The estimates are based on the sample of UK participants. The bars indicate the range of +/-1 standard errors.

identity preferences, as reflected in a higher estimated α.[23]

To explore the identity-threat response prediction, we incorporate measures of both short-term and medium-term team performance. We quantify short-term success using the outcome of a team's most recent match. Specifically, we parameterize the identity tax as:

$$\alpha_{k,m} = \eta \cdot win_{k,m-1} + \delta_k,$$

where $win_{k,m-1}$ is a binary indicator equal to 1 if team k won its previous match, and δ_k captures team-specific factors. Here, η measures how a team's short-run performance influences the identity-incongruence tax, holding constant the team-specific component.

We also examine the medium-term performance by correlating δ_k with the final league ranking from the previous season. Under the identity-threat hypothesis, weaker team performance (or lower league rankings) should correspond to a stronger identity response — that is, a higher α.

Our structural estimation controls for differences in subjective beliefs that might be associated with team performance. This analysis focuses on UK participants, as their sample size per team is large enough for reliable team-specific estimates. In Kenya, the concentration of supporters across only a few teams limits this analysis.

Table 7 reveals that, after controlling for team-specific factors and belief differences, a team winning the previous match reduces the identity-incongruence tax by 9.6 percentage points in the current game. For the medium-term, teams at the bottom of the ranking display systematically higher estimated identity taxes. In our sample, higher-ranked teams exhibit identity taxes ranging from 10% to 25%, while lower-ranked teams show taxes between 30% and 60% (see Figure 5). These findings contradict

[23] A similar pattern would emerge if internal norms against profiting from one's own team's misfortune become more binding when the team underperforms, prompting a stronger identity response in the face of poor results.

Table 7: Heterogeneity in Identity-Incongruence Tax Rate

	PL Ranking	Estimate	Standard Error
η (Change in α if supported team won their previous match)		-0.096***	0.037
δ_k (Team-specific α):			
Manchester City	1	-	-
Manchester United	2	0.122***	0.023
Liverpool	3	0.088***	0.033
Chelsea	4	0.178***	0.038
Leicester City	5	0.245***	0.089
West Ham United	6	0.147	0.187
Tottenham Hotspur	7	0.175***	0.064
Arsenal	8	0.397***	0.028
Leeds United	9	0.018	0.361
Everton	10	0.347***	0.099
Aston Villa	11	0.309*	0.172
Newcastle United	12	0.589***	0.075
Wolverhampton Wanderers	13	0.480***	0.174
Crystal Palace	14	0.405***	0.120
Southampton	15	0.297*	0.153
Brighton and Hove Albion	16	-	-
Burnley	17	0.468***	0.161
Brentford	P	-0.002	0.646
Norwich City	P	0.686***	0.055
Watford	P	0.322	0.318

Notes: This table reports estimates of (1) the change in identity tax if a team won their previous match and (2) the identity-incongruence tax rate for supporters of individual teams. The column "PL Ranking" denotes the team's final position at the end of the Premier League season 2020/21, which directly preceded our study. "P" denotes teams that were promoted from second division to Premier League. The estimates allow for individual-specific beliefs and risk preferences. The estimating equation corresponds to the first-order conditions (equations (10) and (11)), with individual and match-outcome fixed effects accounted for in our estimation. We restricted the sample to teams with at least 20 match-supporter observations and individuals with $r_i > 0$, leading to the exclusion of observations for supporters of Brighton and Hove Albion. Manchester City is also missing because our UK data collection did not include matches involving them. Standard errors are computed via the delta method. The significance levels of estimates are depicted as $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

an opportunistic identity reallocation, where fans would downplay negative outcomes, and support the view that poor performance actually strengthens a source of identity.

Additional details on the team-level decomposition are provided in Table A2. Notably, for lower-ranked teams, we observe not only an increase in α but also a decline in average stakes under neutral beliefs, an increase in both the absolute and relative differences in stakes due to identity, and a smaller share of these differences attributable to identity-specific beliefs.

6.3 Decomposing Identity Investment Differentials

In this section, we use our structural model to decompose the affinity-neutrality investment gap into two components: one driven by identity-distorted beliefs and the other by identity preferences. To illustrate this, we compare the neutral portfolio allocation in equation (6) with the identity-driven allocation in equation (10) for a home win. Formally, we decompose the investment gap for a given match m as follows:

$$\underbrace{x^H_{h,m} - x^N_{h,m}}_{\text{Affinity-driven investment gap}} = \frac{1}{r_m}\left[\underbrace{\Delta P(\bar{\mathbf{p}}^H_m) - \Delta P(\bar{\mathbf{p}}^N_m)}_{\text{Share due to distorted beliefs}} + \underbrace{\text{Ln}\,(1-\alpha) \times \frac{-(\theta_a+\theta_d)}{\Theta}}_{\text{Share due to identity preferences}}\right] \tag{12}$$

Here, $x^H_{h,m}$ and $x^N_{h,m}$ are the budget shares allocated toward a home win in match m by an identity-driven bettor and a neutral bettor, respectively. The terms $\Delta P(\bar{\mathbf{p}}^H_m)$ and $\Delta P(\bar{\mathbf{p}}^N_m)$ capture the odds-adjusted, log-difference in expected marginal returns under supporter beliefs $\bar{\mathbf{p}}^H_m$ and neutral beliefs $\bar{\mathbf{p}}^N_m$, respectively. The term involving α represents the identity tax that discounts the utility of outcomes incongruent with the bettor's identity.

To quantify the contribution of each component, we construct counterfactuals under a common benchmark. We set the odds to $\theta = (3, 3, 3)$, representing fair odds where all outcomes are equally likely.[24] Our counterfactual exercise proceeds as follows:

Beliefs Channel. We first examine how differences in subjective beliefs drive the investment gap. For each match, we generate two counterfactual investment profiles: one using the average beliefs of supporters of the home team or away team, and the other using average beliefs of neutral individuals in the match. We hold risk preferences constant by setting $1/r_m$ equal to the average risk tolerance of neutral supporters in match m.

We then compute the average investment for each scenario. The difference $x(\mathbf{p}^H_m, r) - x(\mathbf{p}^N_m, r)$ captures the effect of identity on investment through the belief channel.[25]

Preference Channel. To isolate the impact of identity preferences, we compute the counterfactual investment using equation (2) for identity-driven bettors with their actual beliefs but with the identity tax α active. In this case, we use the same procedure as above to calculate $x(\mathbf{p}^H_m, r, \alpha) - x(\mathbf{p}^H_m, r)$.

[24] Under this benchmark, if $p = (1/3, 1/3, 1/3)$, a neutral bettor would allocate $x^N_j = 100/3$ tokens to each outcome.

[25] Since ΔP is non-linear in beliefs $\mathbf{p}$, this approach ignores the effect of belief heterogeneity within a match on investment differentials. Rather, it calculates the predicted investment of a 'representative' supporter and a 'representative' neutral in a given match. It preserves, however, the nonlinearity in investments stemming from differences in beliefs across matches.

Because the identity tax enters equation (10) independently of beliefs, the difference attributable to α can be cleanly separated from the belief effect. Because of the non-linear effect of α on investments, we also calculate the predicted investments under heterogeneous α, as estimated in Table 7, for the UK sample. Table 8 presents the decomposition results. For the UK sample (upper panel), we find that identity causes supporters to invest 7.26% of their budget more in their supported team's win. Of this gap, approximately 44% is explained by differences in beliefs between neutrals and supporters, while the remaining 56% arises from identity preferences. Heterogeneous α's increase betting on supported teams winning by 7.09% of the budget. Thus, identity congruence concerns account for 56% to 70% of the affinity-neutrality gap in the UK sample.

For the Kenyan sample, we find slightly larger effects in terms of budget shares, with a 4.72% increase in bets on the supported team winning due to beliefs, and 6.15% due to the identity-incongruence tax α. In relative terms, identity-incongruence concerns account for 57% of the overall differential.

6.4 The Welfare Costs of Identity

In this section, we provide welfare costs of the distortions in investment induced by identity, closely following the decomposition of the affinity-neutral investment gap. We focus on two channels: biased beliefs about match outcomes and the identity-incongruence tax α that discounts payoffs from identity-incongruent outcomes. We evaluate welfare under three normative benchmarks: (i) a policymaker who treats identity preferences as part of true utility but corrects biased beliefs, (ii) a policymaker who also views α as a mistake and removes it holding beliefs fixed, and (iii) the same normative benchmark, recognizing that α amplifies the impact of biased beliefs on portfolio choices. For each case, we translate the estimated distortions into welfare losses as a share of the betting budget and compare magnitudes across the UK and Kenyan samples.

The welfare costs of distorted beliefs. In a first step, we take the perspective of a policymaker who has the same α as the individual, but undistorted probability assessments. We model the distortion in beliefs as a shift Δ of probability from the favored team losing to the favored team winning.[26] The welfare function is defined in equation (17) in appendix B.1.[27] As we show in appendix, the derivative of welfare with regard to the belief distortion is given by

$$\frac{\partial W}{\partial \Delta}/\lambda = \left[\frac{-\Delta}{p_h + \Delta}\frac{\partial x_h}{\partial \Delta} + \frac{\Delta}{p_a - \Delta}\frac{\partial x_a}{\partial \Delta}\right], \tag{13}$$

where $\frac{\partial x_h}{\partial \Delta}$ and $\frac{\partial x_a}{\partial \Delta}$ are the partial derivative of a home-team supporter's investment in a home-team win and home-team loss, respectively, with regard to Δ. This derivative provides an intuition for how much welfare changes when we marginally reduce the belief distortion Δ, in terms of the available budget:[28] equation (13) points out that the impact on welfare is proportional to the sum of relative

[26] This specification leaves the probability of a draw p_d unaffected by the distortion. While simplifying the analysis, this is also broadly consistent with the empirical evidence: As Table 3 shows, supporters perceive the probability of their team winning a game to be significantly higher than neutral individuals, and a nearly identical reduction in the perceived probability of their team losing. This implies that the probability of a draw is similar between supporters and neutrals, consistent with our assumption here.

[27] As for the utility, we could allow an expected utility boost $p_h\eta$ from the team winning. Again, this would not alter any of the portfolio decisions.

[28] The parameter λ converts welfare into budget shares, see appendix B.1 for details.

Table 8: Counterfactual Simulations and Welfare Costs

	Panel A: Counterfactual wager for equal odds across outcomes			
	Neutral Beliefs $x(\bar{p}^N, r)$	Supporter Beliefs $x(\bar{p}^H, r)$	Supporter Beliefs and uniform α $x(\bar{p}^H, r, \alpha)$	Supporter Beliefs and varying α $x(\bar{p}^H, r, \alpha_m)$
UK				
Bet (% of budget)	41.90%	45.05%	49.16%	52.14%
Change in % of budget	-	3.15%	4.11%	7.09%
Kenya				
Bet (% of budget)	49.50%	54.22%	60.37%	-
Change in % of budget	-	4.72%	6.15%	-

	Panel B: Welfare losses in % of budget		
	Belief distortions	Identity-incongruence tax	
		uniform α	match-level α_m
UK			
Policymaker and individuals have same α	0.47%	-	-
Policymaker has $\alpha = 0$	1.25%	0.42%	1.93%
Kenya			
Policymaker and individuals have same α	0.99%	-	-
Policymaker has $\alpha = 0$	2.78%	1.40%	-

Notes: Panel A calculates counterfactual wagers on supported team winning, and Panel B calculates the associated welfare losses. In both panels, the odds for home-team winning, drawing and away-team winning are set to be 1/3. Estimates of match-level α are only available for the UK sample. The three columns simulate the amount of tokens a neutral would bet, a supporter without and with identity-incongruent tax would bet. The first and third rows are calculated based on the equations (6) and (10). They are averaged to match-level first for home-team supporters and away-team supporters separately, before we take the average of the two. We exclude matches with fewer than 20 supporters. Individual-specific r and the corresponding α are applied.

probability distortions, multiplied by the sensitivity of investment with regard to the distortion in beliefs Δ. Setting $p_h = p_a = 0.4$ and $\Delta = 0.04$, as we roughly find in our data, this implies that only 20% of the investment distortion due to beliefs translates into a welfare loss. Hence, a 3% distortion of investment will lead to a welfare loss equivalent to roughly 0.6% of the budget in this calibration.

The example provides an intuition for the magnitudes we should expect. Following the standard approach in behavioral welfare analysis (Bernheim and Taubinsky, 2019; Chetty, 2009), we calculate the welfare loss using a third-order Taylor approximation to the welfare function that takes into account the non-linearity in Δ more accurately. We calculate the welfare losses at the match level, using the same aggregations for risk aversion and beliefs as in the previous section.[29] The results appear in Panel B of Table 8. The welfare loss from distorted beliefs amounts to 0.47% for the UK, and 0.99% for Kenya, reflecting the non-linearity of welfare losses in Δ when identity preferences (including α) are treated as part of true utility.

The welfare costs of the identity-incongruence tax α. In a second step, we abstract from belief distortions and calculate the welfare losses from α, assuming that the policymaker views the identity incongruence tax as a mistake. We hold beliefs fixed in this exercise, to preserve a clear distinction between the welfare costs of beliefs and those of preferences. The welfare function is defined in equation (32) in appendix B.2. Setting $\alpha = 0$ amounts to assuming that the policy maker views betting on game outcomes just like any other financial investment.

As we show in the appendix, the derivative of welfare with regard to the identity incongruence tax α is given by

$$\frac{\partial W}{\partial \alpha}/\lambda = -\frac{\alpha}{(1-\alpha)}\frac{\partial x_h}{\partial \alpha} \tag{14}$$

Similar to the belief distortions before, a one-dollar increase in the bet on a home win only decreases welfare proportional to a factor $1 - \frac{1}{1-\alpha} = -\frac{\alpha}{1-\alpha}$. This is the difference in valuation the policymaker attaches to home-team wins relative to other outcomes (with all being weighted equally) compared to the individual's weight $\frac{1}{1-\alpha}$ relative to the other outcomes. With $\alpha = 0.16$, this amounts to a welfare loss of about 20% of every dollar bet in excess on a home-team win.

To calculate the welfare losses, we again perform a third-order Taylor approximation to account for the non-linearity in α (see section B.2 in the online appendix for details). We follow the same procedure as above to estimate risk aversion and ancillary parameters.

We display the welfare losses in Table 8 for the estimate of the uniform α, as well as the welfare losses using the match-specific estimates of α from Table 7 for the UK. We find that, evaluated at the representative α, welfare losses are 0.4% in the UK sample, and 1.4% in the Kenya sample. The convexity of welfare losses in α leads to higher welfare costs if there is heterogeneity in α compared to the case where we plug in the estimate of the representative α. Taking into account the heterogeneity in α across matches, the welfare losses rise to 1.9% in the UK sample, even though the supporter-weighted mean α is nearly identical to the estimate of the uniform α. These numbers capture the welfare loss

[29] Consistent with how we model belief distortions, we estimate the match-level belief distortion using, both, beliefs of home-team supporters, and away-team supporters relative to neutrals as $\Delta_m = \left(0.5(\bar{p}^H_{h,m} - \bar{p}^N_{h,m}\right) + 0.5(\bar{p}^N_{a,m} - \bar{p}^H_{a,m})\frac{N^H_m}{N^H_m+N^A_m} + \left(0.5(\bar{p}^A_{a,m} - \bar{p}^N_{a,m}\right) + 0.5(\bar{p}^N_{h,m} - \bar{p}^A_{h,m})\frac{N^A_m}{N^H_m+N^A_m}$, where $\bar{p}^K_{k,m}$ indicates the mean belief of supporter type $K \in \{H, A, N\}$ for outcome $k \in \{h, a\}$ in match N. N^K_m is the number of type-K supporters in match m.

from shifting portfolios away from financially optimal bets purely because of identity-driven aversion to "disloyal" gains and highlight the importance of correctly capturing heterogeneity in behavioral biases for welfare (Taubinsky and Rees-Jones, 2017; Allcott and Taubinsky, 2015).

Finally, if the policymaker takes the view that α should not enter the welfare function, this also interacts with the costs of distorted beliefs, as we detail in appendix B.3. Because the distortions in beliefs are added on top of the welfare distortion introduced by α, this raises the welfare losses due to beliefs. Panel B of Table 8 shows that this interaction is quantitatively important: in this benchmark, the welfare loss attributed to belief distortions rises to about 1.3% of the budget in the UK and 2.8% in Kenya. Thus, the welfare costs of identity strongly depend on the policymaker's stance of whether considering α is a preference or a mistake, and belief distortions become especially costly when they operate on top of identity-driven preferences.

Per betting horizon, these identity-driven distortions reduce welfare by roughly 0.5–2.7% of the budget; at the individual level these losses accumulate over repeated betting, and at the market level they scale with the rapid growth of sports betting volumes, making them quantitatively relevant for welfare and regulation. As a back-of-envelope illustration, Americans wagered $121 billion on sports in 2023 (American Gaming Association, T., 2024). If roughly 20 percent of bets involve a supported team, as observed in the data, the implied aggregate welfare loss in the U.S. ranges from approximately $120 million annually—if identity preferences are treated as welfare-relevant and only beliefs are corrected—to $650 million if both belief distortions and heterogeneous identity preferences are treated as mistakes.[30]

7 Discussion

This section discusses policy implications, alternative interpretations, limitations, and applications of the framework.

7.1 Policy Implications for Sports Betting Regulation

Following the 2018 Supreme Court decision that overturned federal restrictions on sports gambling, the rapid expansion of legalized and mobile sports betting has raised concerns about consumer protection (Brown et al., 2025). The evidence presented here suggests that identity-driven biases contribute to excessive betting on favored teams in ways that differ from other forms of gambling. Unlike lotteries, where bettors are not systematically overoptimistic about winnings (Baker et al., 2024; Hollenbeck et al., 2024), sports bettors exhibit both distorted beliefs and identity-driven preferences that sustain unprofitable wagering. The beliefs-versus-preferences decomposition has direct regulatory implications. If belief distortions are the primary channel, disclosure requirements and probability information may reduce the distortion. If identity preferences also play a substantial role, as the estimates suggest, interventions such as modifying default stake allocations or implementing structured cooling-off periods

[30] This calculation assumes that the identity-driven welfare loss rates estimated from Premier League betting apply to sports betting more broadly. The 20 percent figure reflects the share of individual-match observations involving a supported team, which ranges from 12 percent in the UK sample to 28 percent in the Kenyan sample (Table 1). The lower bound applies the 0.5 percent welfare loss (policymaker shares identity preferences, corrects beliefs only) to 20 percent of $121 billion; the upper bound applies the 2.7 percent loss (policymaker treats both beliefs and heterogeneous identity preferences as mistakes) to the same base.

may be necessary to counteract the mechanical shift of budget toward identity-congruent outcomes.

7.2 Concerns and Alternative Interpretations

Functional form of identity preferences. The framework models identity preferences as an identity (disloyalty) tax that is proportional to the utility derived from identity-incongruent outcomes. We justify this choice for two main reasons. First, the proportional tax form introduces a parsimonious yet flexible mechanism to disentangle identity-driven preferences from traditional consumption utility. Its structure aligns with economic settings involving continuous budget allocations, where individuals weigh identity costs against material returns (Akerlof and Kranton, 2000). Second, this form allows for clean empirical identification of identity preferences, as it can be incorporated directly into structural models of decision-making (DellaVigna et al., 2012).

While this functional form is appropriate for our setting, we recognize its limitations. Alternative formulations could include a fixed lump-sum penalty for identity-incongruent actions (similar to participation costs in labor models, Akerlof (1980)) or a non-linear identity tax that grows exponentially with budget share. These extensions would allow for greater flexibility in capturing extreme aversion to disloyalty. However, they also risk complicating estimation, overfitting the data, and diminishing interpretability. Importantly, the proportional structure remains sufficient for modeling the behavior observed in our data while imposing minimal parametric assumptions.

The implications of this choice are clear. The model predicts that individuals' absolute underinvestment in identity-incongruent options will increase in proportion to their financial commitment. This result resonates with theories of identity maintenance (Benjamin et al., 2010) and reflects the intuition and concept grounded in source-dependent utility (Loewenstein and Issacharoff, 1994) and moral contamination models (Stellar and Willer, 2014). The choice of functional form also facilitates counterfactual analyses: calibrated tax rates enable simulations of policies or interventions designed to alleviate identity-driven distortions. To the best of our knowledge, this is the first study to structurally and empirically identify identity preferences as a proportional tax on utility. This innovation advances a growing literature on identity and preferences (Akerlof and Kranton, 2000; Shayo, 2020; Atkin et al., 2021), providing a baseline framework for future work.

Emotional hedging. The emotional hedging hypothesis suggests that fans might bet against their team as a way to cushion the emotional blow if their team loses. Essentially, such bets act as a form of emotional insurance. If emotional hedging is predominant, the prediction would be that supporters allocate less of their investment budget to their own team than neutral supporters. This mechanism predicts a negative identity tax (an identity subsidy). The empirical outcomes do not corroborate the emotional hedging hypothesis as the dominant strategy. We estimate a positive α, suggesting an overarching propensity for identity-consistent investments. This behavior aligns with the overinvestment propensities predicted by the model, specifically for identity-congruent assets (Proposition 1).

Arbitrage opportunities and small-stakes risk aversion. The experiment can present arbitrage opportunities for survey participants. Suppose agents are devoid of identity concerns and are paying attention to betting on matches outside of our experiment. In the case of arbitrage opportunities, they should capitalize on the inflated odds presented in our experiment by wagering on the corresponding outcomes. Consequently, agents should appear risk-neutral as they place their entire budget on the

outcome with the highest odds. However, we find the absence of arbitrage behavior – possibly explained by narrow bracketing, which we do not model explicitly. Such narrow bracketing would also be consistent with risk aversion for small to modest stakes, which otherwise would imply implausible levels of risk aversion for higher stakes (Rabin, 2000). An alternative explanation for this is a noisy mental representation of the true investment opportunities as in Khaw et al. (2021).

Motivated beliefs. Our theoretical framework differentiates the direct impact of identity-specific preferences from the mediating effect of beliefs. However, we do not explicitly model how beliefs are formed. It is conceivable that beliefs arise as a function of identity-specific preferences (motivated beliefs). The literature on motivated beliefs (Bénabou and Tirole, 2016; Zimmermann, 2020) has documented that beliefs, rather than being neutral attempts at accurately predicting possible states of the world, often reflect people's desire for particular states of the world being true. This desire may not only be grounded in economic benefits (Di Tella et al., 2015; Bicchieri et al., 2023), but also, for instance, in ego protection (Drobner and Goerg, 2024), and partisan or group affiliations (Iyengar et al., 2019; Bonomi et al., 2021). Given this potential interrelation, α (identity-incongruence tax) in the model captures the direct impact of identity preferences on choices, controlling for subjective beliefs, asset returns, and risk preferences. However, it does not account for a potential indirect effect of identity preferences on belief formation, i.e. motivated beliefs that are themselves a response to the utility penalty α. Therefore, in our decomposition exercise, the impact of belief distortion on investment choices may reflect both the direct influence of identity on beliefs and the indirect effect of identity preferences on beliefs.

Alternative explanations for the identity-threat response. The finding that the identity-incongruence tax is higher for supporters of underperforming teams is consistent with the identity-threat response, but alternative explanations exist. Supporters might believe in mean reversion, expecting a poorly performing team to improve, and therefore bet more on their team winning out of a belief in the law of small numbers rather than out of identity reinforcement (essentially leading to the gambler's fallacy, as in Rabin (2002)). However, the structural estimation controls for subjective beliefs: the identity-incongruence tax is estimated conditional on reported beliefs, so a supporter who expects mean reversion would express that through higher reported win probabilities, not through a higher α. The identity-threat result therefore reflects residual preference distortions after accounting for any belief-based channels, including mean-reversion expectations.

Sports betting as entertainment. Some people may derive direct consumption utility from sports betting as a form of entertainment. Sports betting is nonetheless best conceived as an investment decision in this context, for the following reasons: First, recent evidence suggests that sports betting is widely used as a substitute for more traditional financial investments. Studying household transaction data in the context of the legalization of sports betting in the US, Baker et al. (2024) find that an additional $1 allocated to online sports betting leads to almost a $1 reduction of investments but has virtually no effect on expenditures for lotteries and other online gambling activities like poker. In a field experiment in Uganda, Herskowitz (2021) finds little evidence for sports betting being a normal consumption good. Instead, individuals use sports betting as a means to finance lumpy expenditures. Second, our study participants are likely to be even less interested in the entertainment aspect of sports betting than the typical sports bettor. People's main motivation for online survey participation is not

entertainment but monetary rewards (Singer and Ye, 2013; Seli et al., 2019; Albert and Smilek, 2023).

If people derive direct consumption utility from sports betting, it is likely to be connected to a live viewing experience. Some people may find it more exciting to watch a game when there is money on the line. How exactly such an entertainment utility would affect the interpretation of our estimates depends on how it interacts with individuals' team allegiances. If the entertainment utility from betting mainly comes from aligning one's money with one's emotional attachment (i.e. the marginal entertainment utility from x^H is higher than from x^N), this would constitute an additional expression of identity preferences, and be captured by our estimate for α. If instead, individuals tend to use monetary bets to add some excitement to neutral games (i.e. the marginal entertainment utility from x^N is higher than from x^H), then our estimate for α would understate the true magnitude of the identity tax.

A related concern is that fans derive utility directly from their team's match outcomes regardless of any bets placed (Kossuth et al., 2020). If supporters experience a utility gain from a win and a utility loss from a defeat, a fully sophisticated bettor might hedge by betting *against* the supported team to smooth consumption across states. The positive estimated α rules out hedging as the dominant motive in the data but does not rule out the presence of direct outcome utility alongside identity-congruent betting. If direct outcome utility is present, the welfare interpretation depends on whether the policymaker views it as part of true utility. Under a policymaker who respects both direct outcome utility and identity preferences, the welfare losses reported here would be lower bounds, as they omit the consumption value of match outcomes themselves. Importantly, two features of the experimental design mitigate the live-viewing channel. First, participants placed bets through an online survey, not at a stadium or while watching a broadcast. Second, bets were placed several days before matches were played, as surveys were fielded early in the week for weekend matches, creating a temporal gap between the betting decision and any potential viewing experience. These features reduce the likelihood that bets were placed primarily to enhance the thrill of watching a match.

7.3 Applications

Consumer Finance. The beliefs-versus-preferences decomposition has direct implications for consumer finance. Investors may exhibit disproportionate favoritism toward domestic equities or bonds for patriotic reasons, yielding an identity tax on portfolio diversity (Huberman, 2015; Foad, 2012). Whether this bias stems from distorted beliefs or genuine preferences has distinct policy implications. If beliefs are the primary channel, informational interventions presenting objective asset comparisons may reduce the distortion. If preferences drive the bias, policymakers may need to design default portfolio allocations that promote diversification without requiring active choice.

Industrial Organization and Marketing. Consumer affiliations may manifest as brand loyalty or preference for products from identity-congruent sources, producing systematic deviations in purchasing behavior analogous to the distortions documented here.[31] The beliefs-versus-preferences decomposition matters for firm strategy: if identity-driven demand reflects distorted beliefs about product quality, informative advertising may neutralize the distortion; if it reflects genuine preferences, product posi-

[31] In a discrete-choice setting, α would enter as a fixed utility tax on identity-incongruent products, while belief distortions would appear as biased perceptions of product quality and/or reliability; the estimation would then rely on standard discrete-choice techniques rather than on the Arrow-Debreu framework.

tioning and pricing are more effective responses.

8 Conclusions

This paper develops and estimates a framework that separately identifies the roles of distorted beliefs and identity-driven preferences in investment decisions. Field experiments with Premier League soccer fans in Kenya and the UK reveal that supporters overinvest in identity-congruent outcomes, with 30–44 percent of the excess investment attributable to belief distortions and the remaining 56–70 percent to an identity-incongruence tax on utility from "disloyal" gains. Our approach offers a methodology to decompose identity-relevant behavior into contributions from identity-specific beliefs and identity-specific preferences. As such, our methodology provides an estimate of the strength of identity concerns that can be compared across groups, settings, and over time. We use this methodology to illustrate its promise for advancing the understanding of identity and its (economic) implications. First, we find that the identity-incongruence tax intensifies when teams underperform, consistent with the identity-threat response. Second, welfare losses from these distortions amount to 0.5–2.7 percent of betting budgets, with the magnitude depending critically on whether identity preferences are treated as part of true utility or as mistakes.

These findings extend beyond sports betting. The beliefs-versus-preferences decomposition applies wherever identity-linked options create a wedge between financially optimal and identity-congruent choices, including domestic versus foreign asset allocation, brand loyalty, and political decision-making. The identity-threat response, documented here in the context of financial payoffs, suggests that identity concerns may intensify precisely when corrective interventions are most needed.

For the regulation of sports betting markets, the key insight is that information provision alone is unlikely to eliminate identity-driven distortions. Because the identity-incongruence tax operates independently of beliefs, interventions that modify choice architecture, such as default stake allocations or cooling-off periods, may be necessary alongside informational approaches.

Future work should explore the decomposition framework in other settings where identity and financial decisions intersect, and investigate whether the identity-threat response operates similarly in domains such as partisan economic behavior and consumer finance.

References

Akerlof, G. A. (1980). A theory of social custom, of which unemployment may be one consequence*. *The Quarterly Journal of Economics*, 94(4):749–775.

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753.

Akerlof, R. (2016). "we thinking"' and its consequences. *American Economic Review*, 106(5):415–419.

Albert, D. A. and Smilek, D. (2023). Comparing attentional disengagement between prolific and mturk samples. *Scientific Reports*, 13(1):20574.

Allcott, H., Boxell, L., Conway, J., Gentzkow, M., Thaler, M., and Yang, D. (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics*, 191:104254.

Allcott, H. and Taubinsky, D. (2015). Evaluating behaviorally motivated policy: experimental evidence from the lightbulb market. *The American Economic Review*, 105(8):2501–2538.

American Gaming Association, T. (2024). Cy 2023 aga commercial gaming revenue tracker. Technical report, American Gaming Association.

Atkin, D., Colson-Sihra, E., and Shayo, M. (2021). How do we choose our identity? a revealed preference approach using food consumption. *Journal of Political Economy*, 129(4).

Baker, L. A. and Emery, R. E. (1993). When every relationship is above average: Perceptions and expectations of divorce at the time of marriage. *Law and human behavior*, 17(4):439–450.

Baker, S. R., Balthrop, J., Johnson, M. J., Kotter, J. D., and Pisciotta, K. (2024). Gambling away stability: Sports betting's impact on vulnerable households. Working Paper.

Bauer, K., Chen, Y., Hett, F., and Kosfeld, M. (2023). Group identity and belief formation: A decomposition of political polarization. *Working Paper*.

Bénabou, R. and Henkel, L. (2025). Identity as self-image. Technical report, National Bureau of Economic Research.

Bénabou, R. and Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30(3):141–64.

Benjamin, D. J., Choi, J. J., and Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, 100(4):1913–28.

Bernheim, B. D. and Taubinsky, D. (2019). Behavioral public economics. In Bernheim, B. D., Della Vigna, S., and Laibson, D. I., editors, *Handbook of Behavioral Economics*, volume 1. Elsevier.

Bicchieri, C., Dimant, E., and Sonderegger, S. (2023). It's not a lie if you believe the norm does not apply: Conditional norm-following and belief distortion. *Games and Economic Behavior*, 138:321–354.

Bohren, J. A., Haggag, K., Imas, A., and Pope, D. G. (2023). Inaccurate statistical discrimination: An identification problem. *Review of Economics and Statistics*, pages 1–45.

Bonomi, G., Gennaioli, N., and Tabellini, G. (2021). Identity, beliefs, and political conflict. *The Quarterly Journal of Economics*, 136(4):2371–2411.

Brown, M., Grasley, N., and Guido, M. (2025). Do sports bettors need consumer protection? evidence from a field experiment. Working Paper.

Butera, L., Metcalfe, R., Morrison, W., and Taubinsky, D. (2022). Measuring the welfare effects of shame and pride. *American Economic Review*, 112(1):122–68.

Cacault, M. P. and Grieder, M. (2019). How group identification distorts beliefs. *Journal of Economic Behavior & Organization*, 164:63–76.

Cassan, G., Keniston, D., and Kleineberg, T. (2021). A division of laborers: Identity and efficiency in india. Technical report, National Bureau of Economic Research.

Charness, G. and Chen, Y. (2020). Social identity, group behavior, and teams. *Annual Review of Economics*, 12:691–713.

Chegere, M., Falco, P., Nieddiu, M., Pandolfi, L., and Stein, M. R. (2022). It's a sure win! experimental evidence on overconfidence in betting behavior. *CSEF WORKING PAPERS*, 655.

Chen, R. and Chen, Y. (2011). The potential of social identity for equilibrium selection. *American Economic Review*, 101(6):2562–2589.

Chen, Y. and Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, 99(1):431–57.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ.*, 1(1):451–488.

Chew, S. H. and Sagi, J. S. (2008). Small worlds: Modeling attitudes toward sources of uncertainty. *Journal of Economic Theory*, 139(1):1–24.

DellaVigna, S., List, J. A., and Malmendier, U. (2012). Testing for altruism and social pressure in charitable giving. *The quarterly journal of economics*, 127(1):1–56.

DellaVigna, S., List, J. A., Malmendier, U., and Rao, G. (2016). Voting to tell others. *The Review of Economic Studies*, 84(1):143–181.

Depetris-Chauvin, E., Durante, R., and Campante, F. (2020). Building nations through shared experiences: Evidence from african football. *American Economic Review*, 110(5):1572–1602.

Deutscher, C., Dimant, E., Humphreys, B. R., and Nowak, A. (2025). Match fixing and sports betting in football: Empirical evidence from the german bundesliga. *Public Finance Review*, 53(5):600–627.

Di Tella, R., Perez-Truglia, R., Babino, A., and Sigman, M. (2015). Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *American Economic Review*, 105(11):3416–3442.

Dimant, E. (2023). Hate trumps love: The impact of political polarization on social preferences. *Management Science*.

Dimant, E., Galeotti, F., and Villeval, M. C. (2024). Motivated information acquisition and social norm formation. *European Economic Review*, 167:104778.

Donkor, K. (2021). The economic value of norm conformity and menu opt-out costs. *Available at SSRN 3955553*.

Drobner, C. and Goerg, S. J. (2024). Motivated belief updating and rationalization of information. *Management Science*, 70(7):4583–4592.

Egan, P. J. and Mullin, M. (2017). Climate change: Us public opinion. *Annual Review of Political Science*, 20:209–227.

Enke, B., Gneezy, U., Hall, B., Martin, D., Nelidov, V., Offerman, T., and Van De Ven, J. (2023). Cognitive biases: Mistakes or missing stakes? *Review of Economics and Statistics*, 105(4):818–832.

Flepp, R., Nüesch, S., and Franck, E. (2016). Does bettor sentiment affect bookmaker pricing? *Journal of Sports Economics*, 17(1):3–11.

Foad, H. S. (2012). Equity home bias and the euro. *The Quarterly Journal of Finance*, 2(01):1250004.

Gadarian, S. K., Goodman, S. W., and Pepinsky, T. B. (2021). Partisanship, health behavior, and policy attitudes in the early stages of the covid-19 pandemic. *Plos one*, 16(4):e0249596.

Gelfand, M. J., Gavrilets, S., and Nunn, N. (2024). Norm dynamics: Interdisciplinary perspectives on social norm emergence, persistence, and change. *Annual Review of Psychology*, 75(1):341–378.

Goette, L., Huffman, D., and Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, 96(2):212–216.

Goette, L., Huffman, D., Meier, S., and Sutter, M. (2012). Competition between organizational groups: Its impact on altruistic and antisocial motivations. *Management Science*, 58(5):948–960.

Grand View Research (2023). Sports betting market size share analysis report, 2030. Accessed: May 24, 2024.

Harrison, G. W. and List, J. A. (2004). Field experiments. *Journal of Economic literature*, 42(4):1009–1055.

Henkel, L. and Zimpelmann, C. (2023). *Proud to not own stocks: How identity shapes financial decisions*. JSTOR.

Herskowitz, S. (2021). Gambling, saving, and lumpy liquidity needs. *American Economic Journal: Applied Economics*, 13(1):72–104.

Hoch, S. J. (1985). Counterfactual reasoning and accuracy in predicting personal events. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4):719.

Hogg, M. A. (1992). *The social psychology of group cohesiveness: From attraction to social identity*. Harvester Wheatsheaf.

Hollenbeck, B., Larsen, P., and Proserpio, D. (2024). The financial consequences of legalized sports gambling. Working Paper.

Huberman, G. (2015). Familiarity breeds investment. *The Review of Financial Studies*, 14(3):659–680.

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., and Westwood, S. J. (2019). The origins and consequences of affective polarization in the united states. *Annual Review of Political cience*, 22:129–146.

Karing, A. (2018). Social signaling and childhood immunization: A field experiment in sierra leone. *University of California, Berkeley Working Paper*.

Kenya National Bureau of Statistics (2022). Economic survey 2022. Technical report, Kenya National Bureau of Statistics.

Khaw, M. W., Li, Z., and Woodford, M. (2021). Cognitive imprecision and small-stakes risk aversion. *The review of economic studies*, 88(4):1979–2013.

Kossuth, L., Powdthavee, N., Harris, D., and Chater, N. (2020). Does it pay to bet on your favourite to win? evidence on experienced utility from the 2018 fifa world cup experiment. *Journal of Economic Behavior & Organization*, 171:35–58.

Kurschilgen, M. (2023). Moral awareness polarizes people's fairness judgments. *Social Choice and Welfare*, 61:339–364.

Kurschilgen, M. and Marcin, I. (2019). Communication is more than information sharing: The role of status-relevant knowledge. *Games and Economic Behavior*, 113:651–672.

Lockwood, B., Allcott, H., Taubinsky, D., and Sial, A. (2021). What drives demand for state-run lotteries? evidence and welfare implications. Technical Report w28975, National Bureau of Economic Research, Cambridge, MA.

Loewenstein, G. and Issacharoff, S. (1994). Source dependence in the valuation of objects. *Journal of Behavioral Decision Making*, 7(3):157–168.

Massey, C., Simmons, J. P., and Armor, D. A. (2011). Hope over experience: Desirability and the persistence of optimism. *Psychological Science*, 22(2):274–281.

Meeuwis, M., Schoar, A., and Simester, D. (2022). Belief disagreement and portfolio choice. *Journal of Finance*, 77(6):3191–3247.

Miller, J. B. and Sanjurjo, A. (2018). Surprised by the hot hand fallacy? a truth in the law of small numbers. *Econometrica*, 86(6):2019–2047.

Morewedge, C. K., Tang, S., and Larrick, R. P. (2018). Betting your favorite to win: Costly reluctance to hedge desired outcomes. *Management Science*, 64(3):997–1014.

Nunn, N. (2022). On the dynamics of human behavior: the past, present, and future of culture, conflict, and cooperation. In *AEA Papers and Proceedings*, volume 112, pages 15–37. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Oh, S. (2023). Does identity affect labor supply? *American Economic Review*, 113(8):2055–2083.

Page, L. (2009). Is there an optimistic bias on betting markets? *Economics Letters*, 102(2):70–72.

Paul, R. J. and Weinbach, A. P. (2009). Sportsbook behavior in the ncaa football betting market: Tests of the traditional and levitt models of sportsbook behavior. *The Journal of Prediction Markets*, 3(2):21–37.

Puri, M. and Robinson, D. T. (2007). Optimism and economic choice. *Journal of financial economics*, 86(1):71–99.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, 68(5):1281–1292.

Rabin, M. (2002). Inference by believers in the law of small numbers. *The Quarterly Journal of Economics*, 117(3):775–816.

Radzevick, J. R. and Moore, D. A. (2008). Myopic biases in competitions. *Organizational Behavior and Human Decision Processes*, 107(2):206–218.

Robinson, J. (2012). Limited insurance within the household: Evidence from a field experiment in kenya. *American Economic Journal: Applied Economics*, 4(4):140–164.

Seli, P., Schacter, D. L., Risko, E. F., and Smilek, D. (2019). Increasing participant motivation reduces rates of intentional and unintentional mind wandering. *Psychological research*, 83(5):1057–1069.

Shayo, M. (2020). Social identity and economic policy. *Annual Review of Economics*, 12:355–389.

Simmons, J. P. and Massey, C. (2012). Is optimism real? *Journal of Experimental Psychology: General*, 141(4):630.

Singer, E. and Ye, C. (2013). The use and effects of incentives in surveys. *The ANNALS of the American Academy of Political and Social Science*, 645(1):112–141.

Statista (2024). Statista. Accessed: May 24, 2024. Available at: https://www.statista.com/outlook/amo/online-gambling/online-sports-betting/united-states.

Stellar, J. E. and Willer, R. (2014). The corruption of value: Negative moral associations diminish the value of money. *Social Psychological and Personality Science*, 5(1):60–66.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, 223:96–103.

Tajfel, H., Billig, M., Bundy, R., and Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1:149–178.

Tajfel, H. and Turner, J. (1979). An integrative theory of intergroup conflict. In Austin, W. G. and Worchel, S., editors, *The Social Psychology of Intergroup Relations*. Wadsworth, Monterey, CA.

Taubinsky, D. and Rees-Jones, A. (2017). Attention variation and welfare: theory and evidence from a tax salience experiment. *The Review of Economic Studies*, 85(4):2462–2496.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, 39(5):806.

Wolfers, J. (2006). Point shaving: Corruption in ncaa basketball. *American Economic Review*, 96(2):279–283.

World Bank (2022). Gdp per capita (current us$) - kenya. https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=KE. Accessed: October 14, 2024.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, 110(2):337–61.

Proof of Proposition 1

Proof: The first-order conditions from equation (1) are

$$u'(\theta_1 x_{i,1}) = \frac{\lambda}{p_{i,1}} \quad \text{and} \quad u'(\theta_1 x_{i,J-1}) = \frac{\lambda}{p_{i,J-1}}$$

We consider two cases for the first-order conditions from equation (2)

Case 1: Beliefs distortion with no identity tax ($\alpha = 0$)

$$u'(\theta_h \hat{x}_{i,h}) = \frac{\hat{\lambda}}{\tilde{p}_{i,h}} \quad \text{and} \quad u'(\theta_h \hat{x}_{i,J-h}) = \frac{\hat{\lambda}}{\tilde{p}_{i,J-h}} \tag{15}$$

Case 2: Beliefs distortion and identity tax ($0 < \alpha < 1$)

$$u'(\theta_h x^*_{i,h}) = \frac{\lambda^*}{\tilde{p}_{i,h}} \quad \text{and} \quad u'(\theta_h x^*_{i,J-h}) = \frac{\lambda^*}{\tilde{p}_{i,J-h}(1-\alpha)} \tag{16}$$

Taking the ratios of the marginal utilities, we have

$$\frac{\frac{u'(\theta_h x^*_{i,h})}{u'(\theta_h \hat{x}_{i,h})}}{\frac{u'(\theta_{J-h} x^*_{i,J-h})}{u'(\theta_{J-h} \hat{x}_{i,J-h})}} = 1 - \alpha$$

Since $0 < 1 - \alpha \leq 1$, and under $u'(\cdot) > 0, u''(\cdot) < 0$, then, $x^*_{i,h} \geq \hat{x}_{i,h}$. QED.

ONLINE APPENDIX

A Structural Model Assuming CRRA Utility

Again, we rely on our model from Section 2 and assume that individuals have CRRA utility, denoted by

$$u(\theta_j x_{i,j}) = \frac{(\theta_j x_{i,j})^{1-\gamma_i}}{1-\gamma_i}, \text{ Where } \gamma_i \neq 1.$$

Here, γ_i signifies the CRRA risk aversion parameter of individual i.

From the first-order conditions (FOCs) neutral bettors we have that

$$x_{i,j}^N = \frac{p_{i,j}\theta_j^{1-\gamma_i}}{p_{i,h}\theta_h^{1-\gamma_i} + p_{i,a}\theta_a^{1-\gamma_i} + p_{i,d}\theta_d^{1-\gamma_i}}.$$

Suppose that individual i supports the home team, then from the FOCs for this non-neutral bettor we have

$$x_{i,h}^H = \frac{\tilde{p}_{i,h}\theta_h^{1-\gamma_i}}{\tilde{p}_{i,h}\theta_h^{1-\gamma_i} + (1-\alpha)\left[\tilde{p}_{i,a}\theta_a^{1-\gamma_i} + \tilde{p}_{i,d}\theta_d^{1-\gamma_i}\right]}$$

$$x_{i,a}^H = \frac{(1-\alpha)\tilde{p}_{i,a}\theta_a^{1-\gamma_i}}{\tilde{p}_{i,h}\theta_h^{1-\gamma_i} + (1-\alpha)\left[\tilde{p}_{i,a}\theta_a^{1-\gamma_i} + \tilde{p}_{i,d}\theta_d^{1-\gamma_i}\right]}.$$

The proof of Proposition 1 under CRRA utility is similar to the proof in Appendix Section 8 above.

B The welfare losses from biased beliefs and α

B.1 The welfare losses from biased beliefs

In this subsection, we characterize the welfare losses from bias in the probability of outcomes. We focus on a particular form of bias, namely that the individual overestimates the probability of the supported team winning by Δ, and underestimates the probability of the other team winning $\tilde{p}_h = p_h + \Delta$ and $\tilde{p}_a = p_a - \Delta$.[32] In this subsection, we abstract from issues arising from α, and assume that it enters the policymaker's objective function

$$W = p_h u(x_h^*(\tilde{\mathbf{p}}, \theta, \alpha)) + p_d(1-\alpha)u(x_d^*(\tilde{\mathbf{p}}, \theta, \alpha)) + p_a(1-\alpha)u(x_a^*(\tilde{\mathbf{p}}, \theta, \alpha)) \tag{17}$$

Following Bernheim and Taubinsky (2019), we start with a given value of Δ, and calculate a third-order Taylor approximation around this value:

$$dW = \frac{\partial W}{\partial \Delta} d\Delta + \frac{1}{2}\frac{\partial^2 W}{\partial \Delta^2}(d\Delta)^2 + \frac{1}{6}\frac{\partial^3 W}{\partial \Delta^3}(d\Delta)^3$$

To obtain the welfare losses from Δ, we then set $d\Delta = -\Delta$.

Proposition 2 (The welfare losses of Δ)**.** *The welfare losses from Δ individuals' decision making for a policymaker with objective function (17) are given by*

$$dW \approx \frac{\lambda \Delta^2}{6}\left[\frac{3p_h^2 + 7p_h\Delta + 6\Delta^2}{(p_h+\Delta)^3}\frac{\partial x_h}{\partial \Delta} - \frac{3p_a^2 - 7p_a\Delta + 6\Delta^2}{(p_a-\Delta)^3}\frac{\partial x_a}{\partial \Delta}\right] > 0 \tag{18}$$

where $\frac{\partial x_h}{\partial \Delta} > 0$ and $\frac{\partial x_a}{\partial \Delta} < 0$ indicate how the investment into the favored (home) team and other (away) team changes as Δ increases, and λ is the Lagrange multiplier on the budget constraint.

The first term in brackets is unambiguously positive. The second term following the minus sign is negative, since $\frac{\partial x_a}{\partial \Delta} < 0$. It can easily be verified that the polynomial in p_a is always positive, irrespective of the value of Δ.[33] This renders the entire term positive. Subsection B.1.1 details the derivation of each of the terms. Notice that α does not appear directly in the expression. This is because α does not directly lead to a distortion according to the policymaker's objective function.[34]

In calculating the welfare losses, we take the following approach: Note that λ is the Lagrange multiplier on the budget constraint of the consumer problem (with Δ present). Hence, its interpretation is the marginal utility of relaxing the budget constraint by one dollar. Thus, overall welfare is approximately equal to $W^* \leq \lambda B$, where B is the budget given to individuals. The inequality results from decreasing marginal utility of money (the infra-marginal units of the budget are worth more than the marginal unit). Thus, by dividing both sides of (18), we can derive an expression for the percent change in welfare due to the removal of bias. This is what we report in Table 8.

We assume CARA utility to calculate the specific values of $\frac{\partial x_h}{\partial \Delta}$ and $\frac{\partial x_a}{\partial \Delta}$ to complete the computations. This is detailed in subsection B.1.2.

[32] The supported team is assumed to be the home team, but this is only to fix notation.

[33] This is easiest to see if we assume the 'worst case' where $\Delta = p_a$. It can easily be seen that the resulting polynomial in p_a is always positive over the range [0,1]

[34] See our discussion below in section B.1.1, specifically equation (21).

B.1.1 Piecing together the terms of the Taylor approximation

We calculate the **first derivative** of the welfare function defined in equation (17) with respect to Δ:

$$\frac{\partial W}{\partial \Delta} = p_h u'_h \cdot \frac{\partial x_h}{\partial \Delta} \cdot \theta_h + p_d u'_d \cdot \frac{\partial x_d}{\partial \Delta} \cdot \theta_d \cdot (1-\alpha) + p_a u'_a \cdot \frac{\partial x_a}{\partial \Delta} \cdot \theta_a \cdot (1-\alpha) \tag{19}$$

Where u'_k denotes the marginal utility of consumption in state k, as defined by equation (16). We substitute the identity $\frac{\partial x_d}{\partial \Delta} = -\frac{\partial x_h}{\partial \Delta} - \frac{\partial x_a}{\partial \Delta}$ and eliminate u'_h, u'_d, u'_a using equations defined by (16). Notice that, because of the latter substitution, all terms involving θ_k's, α, and p_d cancel. We obtain:

$$\frac{\partial W}{\partial \Delta} = \frac{p_h}{p_h + \Delta} \cdot \lambda \cdot \frac{\partial x_h}{\partial \Delta} - \frac{p_d}{p_d} \frac{1-\alpha}{1-\alpha} \cdot \frac{\partial x_h}{\partial \Delta} \cdot \lambda + \frac{p_a}{p_a - \Delta} \cdot \lambda \cdot \frac{\partial x_a}{\partial \Delta} - \frac{\partial x_a}{\partial \Delta} \cdot \lambda \tag{20}$$

Further pulling terms together, we obtain

$$\frac{\partial W}{\partial \Delta} = \left[\frac{-\Delta}{p_h + \Delta} \frac{\partial x_h}{\partial \Delta} + \frac{\Delta}{p_a - \Delta} \frac{\partial x_a}{\partial \Delta} \right] \lambda \tag{21}$$

For the **second and third derivative**, we continue to work with (21), but, following the convention in the literature, we treat the terms $\frac{\partial x_k}{\partial \Delta}$ as constants. We obtain

$$\frac{\partial^2 W}{\partial \Delta^2} = \left[\frac{-p_h}{(p_h + \Delta)^2} \frac{\partial x_h}{\partial \Delta} + \frac{p_a}{(p_a - \Delta)^2} \frac{\partial x_a}{\partial \Delta} \right] \lambda \tag{22}$$

and

$$\frac{\partial^3 W}{\partial \Delta^3} = \left[\frac{2p_h}{(p_h + \Delta)^3} \frac{\partial x_h}{\partial \Delta} + \frac{2p_a}{(p_a - \Delta)^3} \frac{\partial x_a}{\partial \Delta} \right] \lambda \tag{23}$$

Piecing these terms together, we find

$$\frac{dW}{\lambda} = \left(-\frac{\Delta}{p_h + \Delta} \frac{\partial x_h}{\partial \Delta} + \frac{\Delta}{p_a - \Delta} \frac{\partial x_a}{\partial \Delta} \right) d\Delta \tag{24}$$

$$+ \frac{1}{2} \left(-\frac{p_h}{(p_h + \Delta)^2} \frac{\partial x_h}{\partial \Delta} + \frac{p_a}{(p_a - \Delta)^2} \frac{\partial x_a}{\partial \Delta} \right) (d\Delta)^2 \tag{25}$$

$$+ \frac{1}{6} \left(\frac{2p_h}{(p_h + \Delta)^3} \frac{\partial x_h}{\partial \Delta} + \frac{2p_a}{(p_a - \Delta)^3} \frac{\partial x_a}{\partial \Delta} \right) (d\Delta)^3 \tag{26}$$

Substituting $d\Delta = -\Delta$ into these terms yields:

$$\frac{dW}{\lambda} = \left(-\frac{\Delta}{p_h + \Delta} \frac{\partial x_h}{\partial \Delta} + \frac{\Delta}{p_a - \Delta} \frac{\partial x_a}{\partial \Delta} \right) (-\Delta) \tag{27}$$

$$+ \frac{1}{2} \left(-\frac{p_h}{(p_h + \Delta)^2} \frac{\partial x_h}{\partial \Delta} + \frac{p_a}{(p_a - \Delta)^2} \frac{\partial x_a}{\partial \Delta} \right) (-\Delta)^2 \tag{28}$$

$$+ \frac{1}{6} \left(\frac{2p_h}{(p_h + \Delta)^3} \frac{\partial x_h}{\partial \Delta} + \frac{2p_a}{(p_a - \Delta)^3} \frac{\partial x_a}{\partial \Delta} \right) (-\Delta)^3 \tag{29}$$

After some tedious algebra, we arrive at the final expression presented above:

$$dW \approx \frac{\lambda \Delta^2}{6} \left[\frac{3p_h^2 + 7p_h\Delta + 6\Delta^2}{(p_h + \Delta)^3} \frac{\partial x_h}{\partial \Delta} - \frac{3p_a^2 - 7p_a\Delta + 6\Delta^2}{(p_a - \Delta)^3} \frac{\partial x_a}{\partial \Delta} \right] > 0$$

as claimed in Proposition 2.

B.1.2 Calculating the derivatives $\frac{\partial x_h}{\partial \Delta}$ and $\frac{\partial x_a}{\partial \Delta}$

We parametrize the derivatives $\frac{\partial x_h}{\partial \Delta}$ and $\frac{\partial x_a}{\partial \Delta}$ using the optimal choices under a CARA utility function in the main text. Notice that the derivative $\frac{\partial x_k}{\partial \Delta}$ as the total derivative of x_k under the constraint that $dp_h = -dp_a$ and all other parameters remain constant. Thus,

$$\begin{aligned}\frac{\partial x_h}{\partial \Delta} &= \frac{\partial x_h}{\partial p_h} - \frac{\partial x_h}{\partial p_a} \\ &= \frac{1}{\Theta r}\left[(\theta_d + \theta_a)\cdot\frac{1}{p_h} + \theta_d \cdot \frac{1}{p_a}\right] > 0\end{aligned} \tag{30}$$

and

$$\begin{aligned}\frac{\partial x_a}{\partial \Delta} &= \frac{\partial x_a}{\partial p_h} - \frac{\partial x_a}{\partial p_a} \\ &= \frac{-1}{r\Theta}\left[(\theta_d + \theta_h)\frac{1}{p_a} + \theta_d\frac{1}{p_h}\right] < 0\end{aligned} \tag{31}$$

Equations (30) and (31) also verify that the signs of these derivatives are as expected.

B.1.3 Empirical implementation

We proceed in the following steps to calculate the welfare gains in equation (18): first, we need to fix the odds of the games. We set all odds $\theta_k = 3$. This implies that $\Theta = 27$. We take $\Delta = 0.044$ for the UK sample, and $\Delta = 0.07$ for the Kenya sample, based on the results in Table 3.

For each sample (UK and Kenya), we then calculate the average welfare gain according to the following algorithm:[35]

1. For match m, calculate the average beliefs of the neutral subjects $\hat{p}_{h,m} = \frac{1}{K_{N,m}}\sum_{j=1}^{K_{N,m}} p_{h,j,m}$ and $\hat{p}_{a,m} = \frac{1}{K_{N,m}}\sum_{j=1}^{K_{N,m}} p_{a,j,m}$, where $K^{N,m}$ is the number of neutral subjects in match m.
   (a) Given r_i, $\hat{p}_{h,m}, \hat{p}_{a,m}$, calculate $\frac{\partial x_{h,i}}{\partial \Delta}$ and $\frac{\partial x_{a,i}}{\partial \Delta}$ using equations (30) and (31), respectively.
   (b) Given Δ, $\frac{\partial x_{h,i}}{\partial \Delta}$, $\frac{\partial x_{a,i}}{\partial \Delta}$, $\hat{p}_{h,m}, \hat{p}_{a,m}$, calculate $\Delta W_{i,m}$, using equation (18).
2. Given all $\Delta W_{i,m}$, calculate the average $\Delta \bar{W}$ across individuals i and matches m.

B.2 The welfare losses from α

In a second step, we ask what the welfare losses are relative to a benchmark where the policymaker does not think that α belongs into the utility function, but rather is a bias. In this case, the welfare function becomes

$$W = p_h u(x_h^*(\tilde{\mathbf{p}}, \theta, \alpha)) + p_d u(x_d^*(\tilde{\mathbf{p}}, \theta, \alpha)) + p_a u(x_a^*(\tilde{\mathbf{p}}, \theta, \alpha)) \tag{32}$$

[35] Because the welfare gains are not linear in Δ and the beliefs, we cannot simply substitute in means.

While the previous section detailed the calculation of the welfare gains that remained agnostic about the role of α, in this step, we set $\Delta = 0$ and focus on the change in welfare from the viewpoint of the policymaker who considers $\alpha > 0$ a mistake. Starting from α, we calculate

$$dW = \frac{\partial W}{\partial \alpha} d\alpha + \frac{1}{2} \frac{\partial^2 W}{\partial \alpha^2} (d\alpha)^2 + \frac{1}{6} \frac{\partial^3 W}{\partial \alpha^3} (d\alpha)^3$$

and then set $d\alpha = -\alpha$ to calculate the welfare gain from $\alpha = 0$.

Proposition 3 (The welfare losses from α)**.** *The welfare losses from α for a policymaker with the objective function (32) are given by*

$$dW \approx \frac{\alpha^2 \lambda}{(1-\alpha)^2} \frac{\partial x_h}{\partial \alpha} \left[\frac{3 - 7\alpha + 6\alpha^2}{6(1-\alpha)} \right] > 0 \tag{33}$$

where $\frac{\partial x_h}{\partial \alpha} > 0$ is the responsiveness of investment into the favored (home) team winning as α increases, and λ is the Lagrange multiplier on the budget constraint.

Notice, again, that the polynomial in α in the brackets is always positive for any α in [0,1].

We take the same approach in calculating the welfare losses in terms of percent of overall welfare as in subsection B.1, by dividing both sides of equation (33) by $\lambda \cdot B$.

B.2.1 Piecing together the terms of the Taylor approximation

We calculate the **first derivative** of equation (32) with respect to α to obtain

$$\frac{\partial W}{\partial \alpha} = p_h u_h' \Theta_h \frac{\partial x_h}{\partial \alpha} + p_a u_a' \Theta_a \frac{\partial x_a}{\partial \alpha} - p_d u_d' \Theta_d \left[\frac{\partial x_a}{\partial \alpha} + \frac{\partial x_h}{\partial \alpha} \right] \tag{34}$$

$$= \lambda \left[\frac{\partial x_h}{\partial \alpha} + \frac{1}{1-\alpha} \frac{\partial x_a}{\partial \alpha} - \frac{1}{1-\alpha} \frac{\partial x_a}{\partial \alpha} - \frac{1}{1-\alpha} \frac{\partial x_h}{\partial \alpha} \right] \tag{35}$$

$$= \lambda \left[1 - \frac{1}{1-\alpha} \right] \frac{\partial x_h}{\partial \alpha} \tag{36}$$

$$= -\frac{\alpha \lambda}{(1-\alpha)} \frac{\partial x_h}{\partial \alpha} \tag{37}$$

Noting, again, the identity $\frac{\partial x_d}{\partial \alpha} = -\frac{\partial x_h}{\partial \alpha} - \frac{\partial x_a}{\partial \alpha}$, and substituting all the terms u_k' from the first-order conditions (15). Notice that the final term only depends on $\frac{\partial x_h}{\partial \alpha}$. This is intuitive, as α does not induce any distortions in the tradeoff between x_d and x_a, only between x_h and the remainder of the investment.

Continuing as in the previous subsection, we calculate the **second and third derivatives** to be

$$\frac{\partial^2 W}{\partial \alpha^2} = \frac{-\lambda}{(1-\alpha)^2} \frac{\partial x_h}{\partial \alpha} \tag{38}$$

and

$$\frac{\partial^3 W}{\partial \alpha^3} = -\lambda \frac{-2 \cdot (-1)}{(1-\alpha)^3} \frac{\partial x_h}{\partial \alpha} \tag{39}$$

$$= \frac{-2\lambda}{(1-\alpha)^3} \frac{\partial x_h}{\partial \alpha} \tag{40}$$

Piecing all this together and sorting out terms, we obtain

$$dW = \frac{\partial W}{\partial \alpha} d\alpha + \frac{1}{2}\frac{\partial^2 W}{\partial \alpha^2}(d\alpha)^2 + \frac{1}{6}\frac{\partial^3 W}{\partial \alpha^3}(d\alpha)^3 \tag{41}$$

$$= \frac{\alpha^2 \lambda}{1-\alpha}\frac{\partial x_h}{\partial \alpha} \cdot \frac{1-2\alpha}{2(1-\alpha)} + \frac{1}{3}\frac{\lambda \alpha^3}{(1-\alpha)^3}\frac{\partial x_h}{\partial \alpha} \tag{42}$$

$$= \frac{\alpha^2 \lambda}{(1-\alpha)^2}\frac{\partial x_h}{\partial \alpha}\left[\frac{3-7\alpha+6\alpha^2}{6(1-\alpha)}\right] \tag{43}$$

as claimed in Proposition 3.

B.2.2 Calculating the derivative $\frac{\partial x_h}{\partial \alpha}$

To compute $\frac{\partial x_h}{\partial \alpha}$, we again use the CARA assumption. We take the derivative of equation (10) with respect to α to find

$$\frac{\partial x_h}{\partial \alpha} = \frac{1}{(1-\alpha)} \cdot \frac{\theta_a + \theta_d}{r\Theta} > 0 \tag{44}$$

B.2.3 Empirical implementation

We proceed in the following steps to calculate the welfare gains in equation (33): again, we fix the odds of the games by setting all odds $\theta_k = 3$. This implies that $\Theta = 27$. We take $\alpha = 0.166$ for the UK sample, and $\alpha = 0.275$ for the Kenya sample, based on the results in Table 6 (heterogeneous r specifications).

For each sample (UK and Kenya), we then calculate the average welfare gain from removing α according to the following algorithm:[36]

1. For each home-team supporter i in match m, calculate $\frac{\partial x_{h,i}}{\partial \alpha}$ according to equation (44) (they differ in r_i).
2. Calculate the average derivative $\frac{\partial \bar{x}_{h,i}}{\partial \alpha}$ across all individuals i and matches m.
3. Given $\frac{\partial \bar{x}_{h,i}}{\partial \alpha}$ and α, use equation (33) to calculate ΔW.

B.3 The welfare losses from biased beliefs if α doesn't enter the welfare function

The welfare losses from Δ will be larger if α does not enter the policymaker's welfare function, but is present in individual choices. The reason is that α in itself introduces a distortion in investment decisions with which the policymaker disagrees. Adding an additional distortion in beliefs consequently increases the associated welfare losses. This modifies the results from appendix B as follows:

Proposition 4 (The welfare losses of Δ when α doesn't enter the objective function of the policymaker)**.** *The welfare losses from Δ individuals' decision making for a policymaker with objective function (32) are given by*

$$dW \approx \frac{\lambda \Delta^2}{6}\left[\frac{3p_h^2 + 7p_h\Delta + 6\Delta^2}{(p_h+\Delta)^3}\frac{\partial x_h}{\partial \Delta} - \frac{1}{1-\alpha}\frac{3p_a^2 - 7p_a\Delta + 6\Delta^2}{(p_a-\Delta)^3}\frac{\partial x_a}{\partial \Delta}\right] + \frac{\alpha\lambda\Delta}{1-\alpha}\frac{\partial x_h}{\partial \Delta} > 0 \tag{45}$$

[36] Because the welfare gains are not linear in the beliefs, we cannot simply substitute in means.

where $\frac{\partial x_h}{\partial \Delta} > 0$ *and* $\frac{\partial x_a}{\partial \Delta} < 0$ *indicate how the investment into the favored (home) team and other (away) team changes as* Δ *increases, the parameter* α *is the identity incongruence tax,* λ *is the Lagrange multiplier on the budget constraint.*

B.3.1 Piecing together the terms of the Taylor approximation

We calculate the **first derivative** of the welfare function defined in equation (32) with respect to Δ:

$$\frac{\partial W}{\partial \Delta} = p_h u'_h \cdot \frac{\partial x_h}{\partial \Delta} \cdot \theta_h + p_d u'_d \cdot \frac{\partial x_d}{\partial \Delta} \cdot \theta_d + p_a u'_a \cdot \frac{\partial x_a}{\partial \Delta} \cdot \theta_a \tag{46}$$

Where u'_k denotes the marginal utility of consumption in state k, as defined by equation (16). We substitute the identity $\frac{\partial x_d}{\partial \Delta} = -\frac{\partial x_h}{\partial \Delta} - \frac{\partial x_a}{\partial \Delta}$ and eliminate u'_h, u'_d, u'_a using equations defined by (16). Notice that, because of the latter substitution, all terms involving θ_k's and p_d cancel, as before. However, some terms involving α remain. Proceeding as in section B.1, we obtain:

$$\frac{\partial W}{\partial \Delta} = \frac{p_h}{p_h + \Delta} \cdot \lambda \cdot \frac{\partial x_h}{\partial \Delta} - \frac{p_d}{p_d} \frac{1}{1-\alpha} \cdot \frac{\partial x_h}{\partial \Delta} \cdot \lambda + \frac{\lambda}{1-\alpha} \cdot \left(\frac{p_a}{p_a - \Delta} \frac{\partial x_a}{\partial \Delta} - \frac{\partial x_a}{\partial \Delta} \right) \tag{47}$$

where the terms $\frac{1}{1-\alpha}\cdot$ magnify the corresponding components (related to changes in investments x_d and x_a) that the policymaker views as additionally distorted by α. Noticing that $\frac{1}{1-\alpha}\cdot = 1 + \frac{\alpha}{1-\alpha}$, we can rewrite equation (47) as

$$\frac{\partial W}{\partial \Delta} = \left[\frac{-\Delta}{p_h + \Delta} \frac{\partial x_h}{\partial \Delta} + \frac{1}{1-\alpha} \frac{\Delta}{p_a - \Delta} \frac{\partial x_a}{\partial \Delta} \right] \lambda - \frac{\alpha \lambda}{1-\alpha} \frac{\partial x_h}{\partial \Delta} \tag{48}$$

Thus, removing α from the objective function of the policy maker changes the welfare calculations in two ways: First, it scales up terms involving $\frac{\partial x_a}{\partial \Delta}$ by $1/(1-\alpha)$, reflecting the additional welfare losses that changes in beliefs have on the relative allocation between x_a and x_h introduced by α. Second, it adds an additional term to the first derivative involving only α and $\frac{\partial x_h}{\partial \Delta}$, reflecting the welfare costs in the overinvestment into x_h. Notice that (2) will have no effect on higher-order derivatives, as Δ does not appear in it, whereas (1) simply leads to a rescaling of some of the terms in higher-order derivatives.

Carrying these two changes forward through the rest of the derivation from appendix B, we obtain

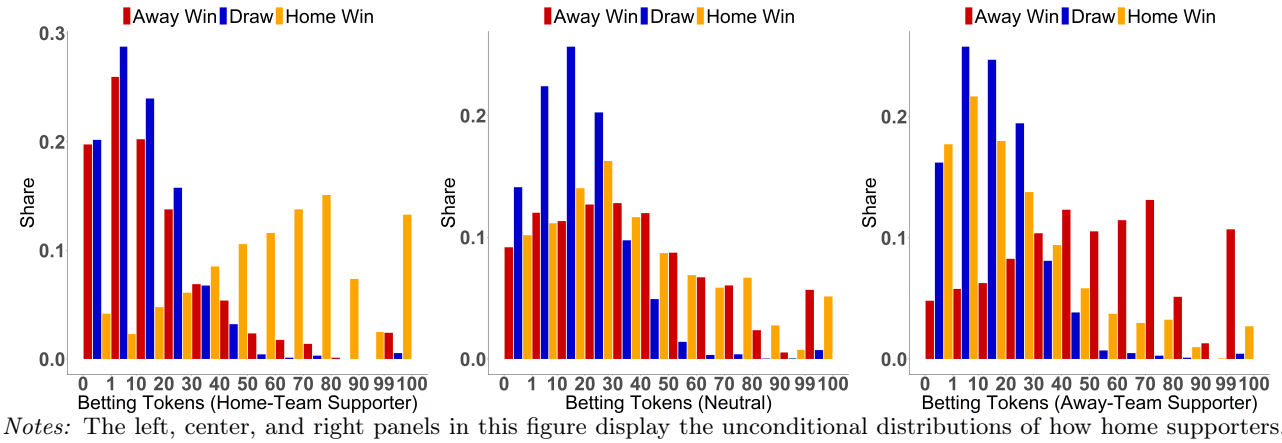
$$dW \approx \frac{\lambda \Delta^2}{6} \left[\frac{3p_h^2 + 7p_h\Delta + 6\Delta^2}{(p_h + \Delta)^3} \frac{\partial x_h}{\partial \Delta} - \frac{1}{1-\alpha} \frac{3p_a^2 - 7p_a\Delta + 6\Delta^2}{(p_a - \Delta)^3} \frac{\partial x_a}{\partial \Delta} \right] + \frac{\alpha \lambda \Delta}{1-\alpha} \frac{\partial x_h}{\partial \Delta} > 0 \tag{49}$$

The empirical implementation is the same as in appendix B, but adding the estimate of the homogenous α, in parallel to appendix B.2.

C Supplementary tables and figures

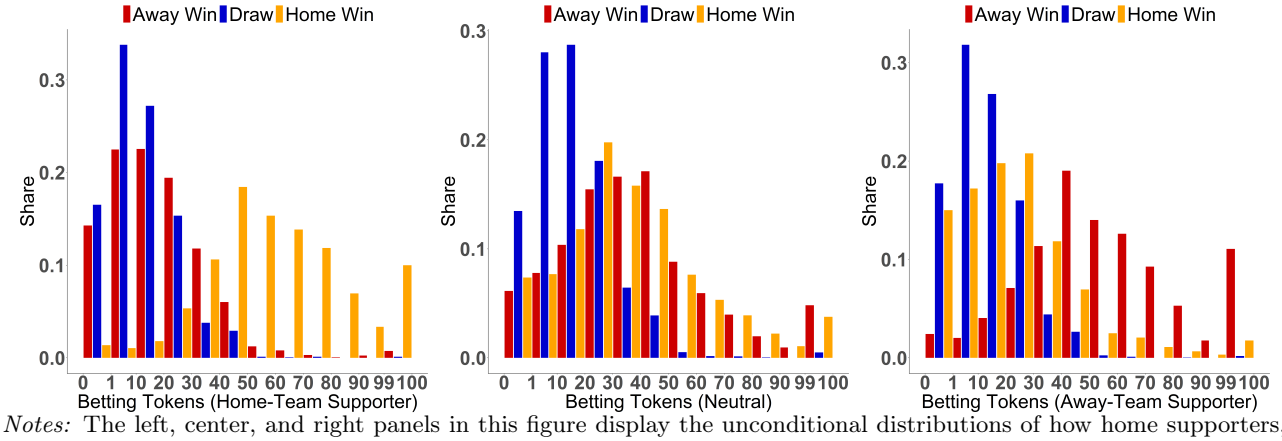
C.1 Figures

Figure A1: Portfolio Allocations (UK): Unconditional Distributions



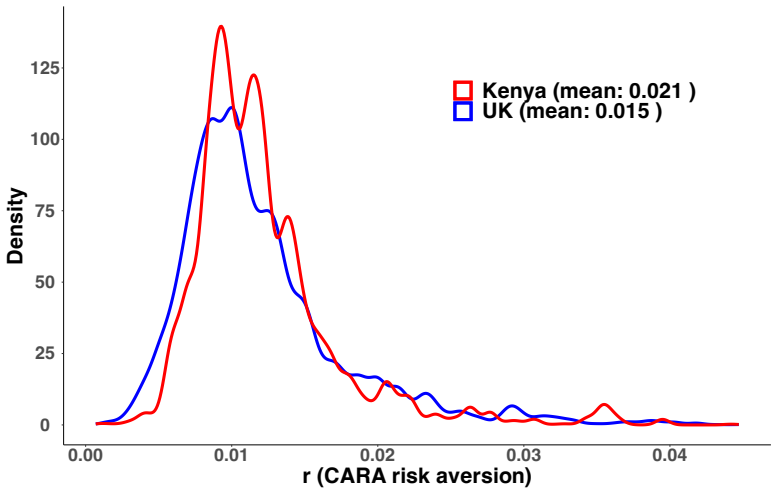
Notes: The left, center, and right panels in this figure display the unconditional distributions of how home supporters, neutrals, and away supporters in the UK allocate their betting budget across the three potential outcomes of a soccer match (home-win in orange, draw in blue, away-win in red).

Figure A2: Portfolio Allocations (Kenya): Unconditional Distributions



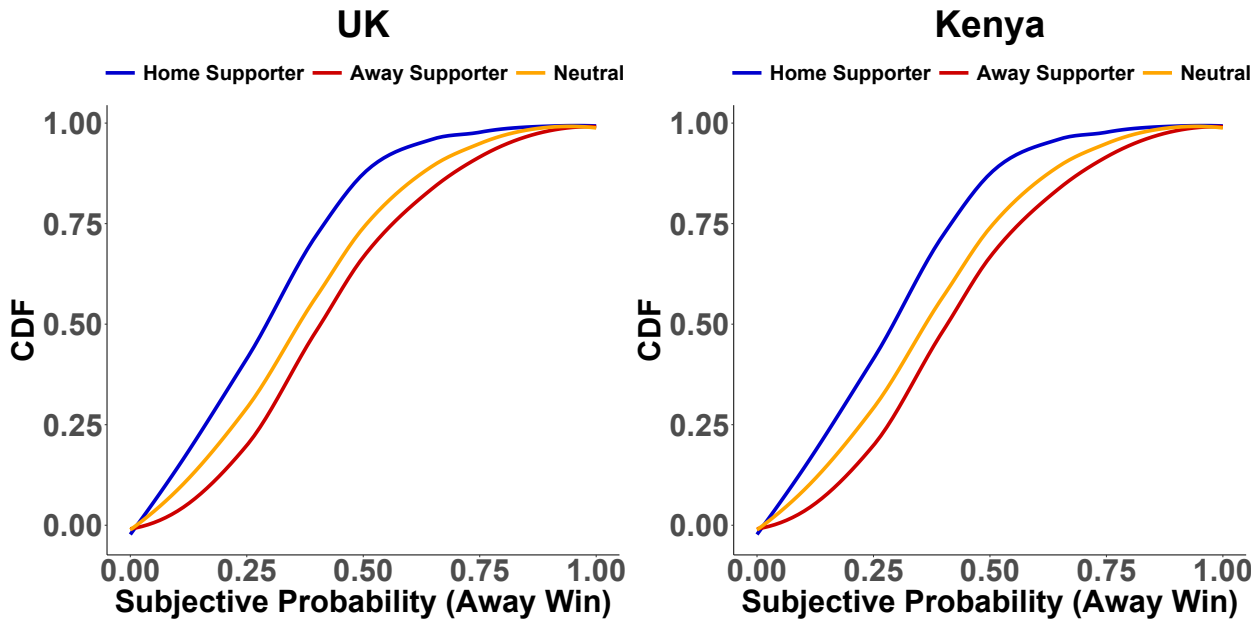
Notes: The left, center, and right panels in this figure display the unconditional distributions of how home supporters, neutrals, and away supporters in Kenya allocate their betting budget across the three potential outcomes of a soccer match (home-win in orange, draw in blue, away-win in red).

Figure A3: Distribution of Risk Preferences



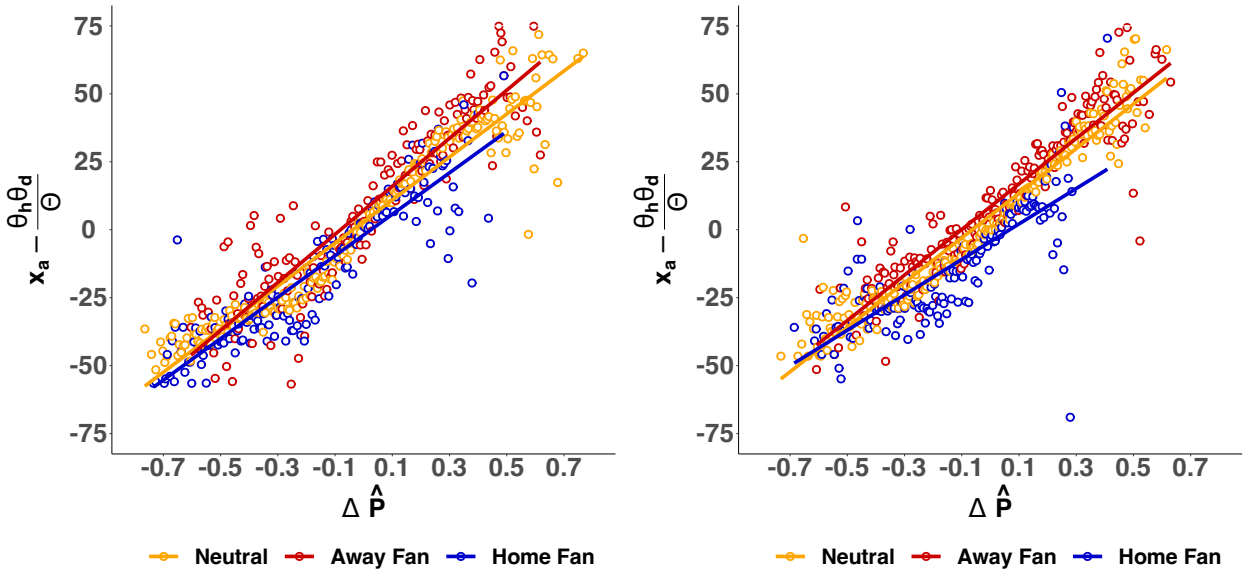
Notes: This figure presents the distribution of individual risk aversion estimates under the assumption of CARA and based on equation (8), separately for participants in the UK (blue) and Kenya (red).

Figure A4: Subjective Beliefs (CDFs): Away-Team Win



Notes: The left and right panels above show the CDFs of survey participants' self-reported beliefs about "away team winning" of Premier League soccer matches for the UK and Kenya, respectively. CDFs are first calculated for each match and then averaged across all matches. The CDFs in each panel are categorized by whether the participants support the home or the away team or are neutral across the teams playing.

Figure A5: Semi-structural Portfolio Allocation (Conditional on Beliefs)



Notes: The panels in the figure above show a scatter plot of actual wagers on away-win against the model-computed first-order conditions of neutral bets on away-win for the UK and Kenya, respectively. The points in each panel are color-coded by supporter status: yellow for neutral bettors, blue for home-team supporters, and red for away-team supporters. The vertical axis is the wager on away-win minus the ratio of odds. The horizontal axis is ΔP and captures the non-linear relationship of how beliefs and odds affect the amount wagered on a potential outcome.

C.2 Tables

Table A1: Shares (%) of Supporters by Team

	Kenya Sample	UK Sample
Manchester United	35.64	23.94
Liverpool	4.28	20.71
Arsenal	17.63	11.50
Chelsea	31.11	9.39
Tottenham Hotspur	0.63	8.08
Manchester City	8.94	5.41
Newcastle United	0.13	4.60
Westham United	0.13	2.86
Everton	0	2.67
Leicester City	0.25	2.24
Southampton	0	1.49
Wolverhampton Wanderers	0.13	1.43
Brighton and Hove Albion	0	1.31
Crystal Palace	0.25	0.93
Leeds United	0	0.87
Aston Villa	0.13	0.81
Burnley	0.25	0.68
Watford	0.13	0.67
Norwich City	0	0.5
Brentford	0.25	0.19

Notes: In this table, we portray the share of supporters for specific teams for both Kenya (column (1)) and the UK (column (2)), with teams sorted by their share of supporters in the UK sample.

Table A2: Counterfactual Simulation at Team Level (UK Sample)

Team	Rank	α	Neutral Beliefs $\bar{x}^N(p,r)$	Supporter Beliefs $\bar{x}^N(\tilde{p},r)$	Supporter Beliefs + Identity $\bar{x}^H(\tilde{p},r)$	N	$\%\Delta_1$	$\%\Delta_2$
ManCity	1	-	-	-	-	-	-	-
ManUnited	2	0.12	53.45	57.56	60.66	639.5	7.69	13.49
Liverpool	3	0.09	60.65	66.26	68.27	295.0	9.25	12.57
Chelsea	4	0.18	63.74	65.39	69.70	225.0	2.59	9.35
Leicester	5	0.24	32.59	37.99	43.49	40.0	16.57	33.43
West Ham	6	0.15	22.74	23.10	26.75	21.0	1.58	17.64
Spurs	7	0.17	50.21	55.50	59.66	171.0	10.55	18.83
Arsenal	8	0.40	48.62	53.87	63.87	311.0	10.80	31.35
Leeds United	9	0.02	20.24	23.85	24.19	5.5	17.86	19.55
Everton	10	0.35	27.96	28.07	37.87	39.0	0.39	35.44
Aston Villa	11	0.31	31.44	31.23	38.68	10.0	-0.65	23.05
Newcastle	12	0.59	17.12	17.68	38.98	30.0	3.25	127.68
Wolvers	13	0.48	18.85	22.48	35.54	13.0	19.28	88.59
Crystal Palace	14	0.40	30.92	34.25	45.24	10.5	10.78	46.30
Southampton	15	0.30	22.26	30.23	38.19	11.0	35.79	71.57
Brighton	16	-	20.82	25.52	-	15.0	22.59	-
Burnley	17	0.47	18.71	29.75	42.51	7.0	59.04	127.29
Brentford	-	-0.00	-	-	-	-	-	-
Watford	-	0.32	20.79	43.53	49.59	6.0	109.39	138.50
Norwich City	-	0.69	17.42	-0.37	21.60	4.0	-102.10	24.02
Average	-	$\alpha>0$	32.14	35.88	44.99	103.0	11.65	40.00
Avg. (N>20)	-	$\alpha>0$	41.90	45.05	52.14	218.8	7.52	24.44

Notes: This table simulates counterfactual wagers on team winning. The odds for home-team winning, drawing and away-team winning are set to be equal. Columns (4), (5) and (6) simulate the amount of tokens a neutral would bet, a supporter without and with identity-incongruent tax would bet. They are calculated based on the equations (6) and (10). They are averaged to team-level first for home-team supporters and away-team supporters separately, and then averaged across the two. Heterogeneous r and corresponding α are applied. Column (7) is the number of supporters for each team. Columns (8) and (9) show the percentage changes $\%\Delta_1 = (\bar{x}^N(\tilde{p},r) - \bar{x}^N(p,r))/\bar{x}^N(p,r)$ and $\%\Delta_2 = (\bar{x}^H(\tilde{p},r) - \bar{x}^N(p,r))/\bar{x}^N(p,r)$, respectively. The simulation is calculated only for the UK.

Table A3: Semi-structural Portfolio Allocation (Heterogeneous Risk Preferences & Bookmaker Beliefs)

	Dependent Variable: $x - \frac{\theta_d \theta_j}{\Theta} - r_i^{-1} \Delta P^{BM}$			
	UK Home Win	UK Away Win	Kenya Home Win	Kenya Away Win
Home-Team Supporter	7.038^{***}	-5.369^{***}	6.766^{***}	-6.888^{***}
	(0.773)	(0.707)	(0.613)	(0.608)
Away-Team Supporter	-5.667^{***}	6.721^{***}	-6.566^{***}	6.879^{***}
	(0.667)	(0.712)	(0.591)	(0.608)
Individual FE	Yes	Yes	Yes	Yes
Match FE	Yes	Yes	Yes	Yes
Observations	19,639	19,639	18,408	18,408
R^2	0.209	0.214	0.178	0.174

Notes: This table presents semi-structural estimates corresponding to the first-order conditions for neutral bets computed separately for the UK (columns (1) and (2)) and Kenya (columns (3) and (4)). In each column, we regress wagers on home-win (or away-win) x_j minus $\frac{\theta_d \theta_j}{\Theta} - r^{-1} \Delta P^{BM}$ on indicators for home- and away-team supporters. In this table, we assume that individuals use bookmaker beliefs when betting. Therefore, r_i^{-1} and ΔP^{BM} are estimated at the individual level where the corresponding bookmaker beliefs replace subjective beliefs. We exclude individuals for whom we estimate zero or negative levels of risk aversion (see section 5.2). We cluster our standard errors at the individual level and report the robust standard errors for our estimates in parentheses. The significance levels of our estimates are depicted as $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C.3 Details about calculating randomized odds

For each match, we calculated 7 sets of home and away odds θ_h^z, θ_a^z that would yield wedges $z \in \{-2.5, -1.5, -.5, 0, +.5, +1.5, +2.5\}$,[37] where

$$z = p_a^0 * \theta_a^z - p_h^0 * \theta_h^z \tag{50}$$

To guarantee that the sets of odds would keep constant the overall assessed likelihood of match outcomes (as implied by bookmakers' odds) $\frac{1}{\theta_h^z} + \frac{1}{\theta_d^z} + \frac{1}{\theta_a^z} = \frac{1}{\theta_h^0} + \frac{1}{\theta_d^0} + \frac{1}{\theta_a^0}$, the sets of odds also had to satisfy the condition $\theta_a^z = \frac{1}{p_a^0 - x^z}$, $\theta_h^z = \frac{1}{p_h^0 + x^z}$. This way, we ensured that an increase in the odds for a loss that implied e.g., a 5 percentage point drop in likelihood was associated with a decrease in odds for win that implied a 5 percentage point increase in likelihood. Plugging this back into equation (50) above, we get

$$p_a^0 * \frac{1}{p_a^0 - x^z} - p_h^0 * \frac{1}{p_h^0 + x^z} = z \tag{51}$$

or in terms of bookmakers' benchmark odds:

$$\frac{1}{\theta_a^0} * \frac{1}{\frac{1}{\theta_a^0} - x^z} - \frac{1}{\theta_h^0} * \frac{1}{\frac{1}{\theta_h^0} + x^z} = z \tag{52}$$

Solving for x^z yields, for positive wedges z:

$$x = -\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z} + \sqrt{(\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z})^2 + \frac{1}{\theta_a^0\theta_h^0}} \tag{53}$$

For z taking negative values, x^z is given by:

$$x = -\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z} - \sqrt{(\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z})^2 + \frac{1}{\theta_a^0\theta_h^0}} \tag{54}$$

[37] The wedge is 0 for the benchmark odds.

Table A4: Examples of Randomized Odds θ_h^z, θ_a^z for a Set of Benchmark Odds and Wedges z

	Benchmark odds θ_h^0, θ_a^0:						
Wedge	1.1 (θ_h^0), 11 (θ_a^0)	1.2, 8.0	1.5, 5.0	2.0, 2.0	5.0, 1.5	8.0, 1.2	11.0, 1.1
2.5	1.03, 37.77	1.09, 27.23	1.24, 16.63	1.19, 6.19	1.59, 4.23	1.54, 3.23	1.53, 2.90
(x=...)	x=.064	x=.088	x=.140	x=.339	x=.43	x=.524	x=.56
1.5	1.04, 26.89	1.10, 19.35	1.28, 11.76	1.30, 4.30	1.95, 2.83	2.06, 2.11	2.16, 1.87
(x=...)	x=.054	x=.073	x=.115	x=.268	x=.314	x=.359	x=.373
.5	1.07, =16.16	1.15, 11.64	1.38, 7.09	1.62, 2.62	3.36, 1.76	4.86, 1.33	6.36, 1.19
(x=...)	x=.029	x=.039	x=.059	x=.118	x=.098	x=.081	x=.066
0	1.1, 11	1.2, 8.0	1.5, 5.0	2.0, 2.0	5.0, 1.5	8.0, 1.2	11.0, 1.1
(x=...)	x=0	x=0	x=0	x=0	x=0	x=0	x=0
-.5	1.19, 6.36	1.33, 4.86	1.76, 3.36	2.62, 1.62	7.09, 1.38	11.64, 1.15	16.16, 1.07
(x=...)	x=-.066	x=-.081	x=-.098	x=-.118	x=-.059	x=-.039	x=-.029
-1.5	1.87, 2.16	2.11, 2.06	2.83, 1.95	4.30, 1.30	11.76, 1.28	19.35, 1.10	26.89, 1.04
(x=...)	x=-.373	x=-.359	x=-.314	x=-.268	x=-.115	x=-.073	x=-.054
-2.5	2.90, 1.53	3.23, 1.54	4.23, 1.59	6.19, 1.19	16.63, 1.24	27.23, 1.09	37.77, 1.03
(x=...)	x=-.56	x=-.524	x=-.43	x=-.339	x=-.140	x=-.088	x=-.064

Notes: In this table, we show for different sets of benchmark odds θ_h^0 and θ_a^0 which odds participants would be shown if they were randomized according to wedges -2.5, -1.5, -.5, 0, .5, 1.5, or 2.5. The x-term indicates the underlying difference in assessed probabilities between the benchmark odds and the randomized odds.