The most recent version of this working paper can always be downloaded by following: this link.

Identity and Economic Incentives

Kwabena Donkor* Lorenz Goette† Maximilian W. Müller‡

Eugen Dimant§ Michael Kurschilgen¶

December 20, 2023

Abstract

This paper examines how beliefs and preferences drive identity-conforming consumption or investments. We introduce a theory that explains how identity distorts individuals' beliefs about potential outcomes and imposes psychic costs on benefiting from identity-incongruent sources. We substantiate our theoretical foundation through two lab-in-field experiments on soccer betting in Kenya and the UK, where participants either had established affiliations with the teams involved or assumed a neutral stance. The results indicate that soccer fans have overoptimistic beliefs about match outcomes that align with their identity and bet significantly higher amounts on those than on outcomes of comparable games where they are neutral. After accounting for individuals' beliefs and risk preferences, our structural estimates reveal that participants undervalue gains from identity-incongruent assets by 9% to 27%. Our counterfactual simulations imply that identity-specific beliefs account for 30% to 44% of the investment differences between neutral observers and supporters, with the remainder being due to identity preferences.

*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).

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, UMich, NUS, Northwestern Kellogg, MIT Sloan, Cornell, QME, Lausanne, Fribourg, Uppsala, Frankfurt, and briq 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.

1 Introduction

Research in psychology, sociology, and economics has shown that social identity significantly impacts individuals' beliefs, potentially leading to biased perceptions and decisions. These affiliations predispose individuals to avoid actions misaligned with their identity, even at a financial cost, or can trigger sanctions or psychological distress for non-conforming behaviors (Akerlof and Kranton, 2000). This phenomenon introduces a challenge: distinguishing between the influences of beliefs and preferences on choices.[1] While traditional economic models simplify this complexity through assumptions like rational expectations and stable preferences, they may neglect the intricate relationship between social identity and decision-making. Given the growing polarization across numerous domains such as politics, environmental issues, and health, unraveling the underlying root causes of behavior—beliefs or preferences—is crucial for understanding behavior and improving policy design.[2]

This paper explores the impact of social identity on economic decision-making by quantifying the influence of identity-specific beliefs and preferences. We introduce a model wherein identity distorts individuals' beliefs about potential outcomes and inflicts psychic costs for gains derived from identity-incongruent sources. We substantiate our theoretical foundation through two large-scale field experiments on soccer betting in Kenya and the UK, where participants either had established affiliations with the teams involved or assumed a neutral stance. Finally, we integrate the model and the experiments within a structural analysis to measure and distinguish between the importance of identity-specific beliefs and preferences for observed choices.

Our proposed theory of investment decision-making focuses on how agents' affiliations and identification with specific assets shape their portfolio choices. Agents in the model exhibit affinity towards certain assets, a phenomenon we refer to as *identity concerns*. This affinity creates complementarities between agents' identities and their investment choices. These associations may manifest in various forms, such as preferential stock holdings in an employer's company (Meulbroek, 2005), investments in personal ventures, or patriotic allocations to domestic firms (Foad, 2012). Affiliations may also include investments driven by environmental motivations, as seen in sustainability or green energy investing (Zerbib, 2022), or those guided by social cause affiliations aligning with an investor's values or ideologies.

The model builds on the traditional subjective expected utility maximization framework

[1] A key challenge in understanding these dynamics and identifying the underlying mechanisms is differentiating the effects of beliefs and preferences on decision-making–a task complicated by multiple belief-preference combinations potentially explaining a single choice pattern (Manski, 2004).

[2] Consider for example areas like discrimination, where understanding the underlying motivations – be they taste-based or belief-based—carries significant policy and welfare implications (Bohren et al., 2023).

augmented with identity concerns entering through distortions in beliefs and distaste for benefiting from identity-incongruent outcomes. The model's main prediction is that individuals overinvest in identity-congruent assets compared to neutral assets after accounting for subjective beliefs, asset returns, and risk preferences. We test this prediction in two lab-in-the-field experiments on soccer betting among English Premier League (PL) fans in the UK and Kenya.

Soccer betting offers an ideal empirical context characterized by strong team affiliations (fan identity) and the necessity to balance monetary gains against identity concerns. We recruited 802 soccer fans from Kenya and 1608 from the UK for our experiment before the 2021/22 Premier League season. During the season, we tasked participants with placing bets for upcoming matchdays. For each matchday, we privately elicited fans' subjective beliefs about each game's potential outcomes to directly measure subjective expectations—following Manski (2004). Then, we had fans allocate a given investment budget across these outcomes, where we experimentally varied the payoffs for outcomes. After each matchday, we paid participants based on their wagers on one randomly selected game per matchday.

Key features of our experimental design are (1) we maintain a consistent decision problem across matches, irrespective of whether the participant supports a team or remains neutral, (2) we exogenously vary asset returns for the same match and across participants. Using bookmakers' predictions as proxies for objective benchmarks of forecasts of match outcomes, we can compare fans' subjective beliefs against these benchmarks and assess distortions in beliefs related to fan identity.[3] Exogenously varying asset returns, we create differing tradeoffs that allow us to examine whether participants' earnings maximization behavior when they support a team deviates from when they maintain neutrality.

Our reduced-form analysis reveals that UK and Kenyan fans allocate more of their budgets to bets on their favored teams than when they are neutral. This behavior results in a 20% higher budget allocation to identity-relevant assets than neutral assets (*neutral-supporter gap*). Furthermore, while neutral fans' predictions generally align with bookmakers' forecasts, supporters believe their team's chances of winning are 10% to 18% higher than assessed by neutral fans. These disparities in budget allocation and beliefs hold even when we account for individual and match-specific fixed effects, suggesting a deep-rooted influence of identity on fans' perceptions and decision-making processes.

To explore whether differences in risk preference might explain the neutral-supporter gap, we conduct a semi-structural analysis incorporating individual risk preferences. The results

[3] Having a proxy for an objective likelihood benchmark represents a significant advantage of our experimental context, a feature often hard to replicate in other scenarios, such as the inability to vote neutrally in elections or the difficulty in accurately predicting market reactions to unforeseen global events.

refute the notion that bettors, regardless of team loyalty, mimic neutral bettors' portfolio allocations after adjusting for subjective beliefs, asset returns, and risk preferences. Instead, we find that team loyalty plays a pivotal role.

In our structural model, we quantify the neutral-supporter gap by introducing an "*identity-incongruence tax*", meaning supporters assign less value to the benefits from earnings when their team does not win. Consequently, the marginal utility of a dollar earned when one's team does not win is lower than from a win. We estimate identity-incongruence tax rates of 17% for the UK and 27% for Kenya. Delving deeper into the heterogeneity of these taxes, we observe a notable correlation with teams' performance—both with a medium-run measure, a team's final league ranking from the previous season, and with a short-run measure, a win in the previous game. Specifically, our estimated tax rates range from 12% for the top-ranked team in our sample to 47% for the lowest-ranked team, a fourfold increase. Similarly, we find that the identity tax is ten percentage points lower for supporters of teams with a recent win. The finding that a team's poor performance might intensify supporters' identity preferences resonates with the notion of an increased self-signaling value of betting on the team's success despite its poor performance, and with the idea that internal norms against benefiting from one's team's misfortune are stronger when the team is struggling.

Finally, simulating counterfactuals with our model parameters, we find that identity-specific beliefs account for one-third to at most half of the investment gap between neutral observers and supporters, with the rest attributed to identity preferences. Supporters in the UK and Kenya have 21% to 24% higher investment in their team winning than neutral fans' investment. For Kenya, our estimates suggest almost equal contributions from belief differences and identity preferences for the observed neutral-supporter investment gap. In the UK, on the other hand, only one-third is due to belief differences, with the remaining driven by identity preferences.

Our model and findings that identity influences behavior through both beliefs and preferences offer crucial insights for policy design and welfare economics across various contexts. For example, our framework applied to consumer finance can explain tendencies like patriotic investment in domestic firms (Foad, 2012) or employees' preference for their employer's stocks in pension plans (Meulbroek, 2005). Potential ways of addressing these biases are the need for diversified financial products and educational programs that consider identity concerns. Our framework also provides insights into firms' competitive strategy, e.g., to address consumer biases due to brand loyalty or identity affiliations (Park and Srinivasan, 1994; Butler, 2018). Firms could use informative advertising if such consumer biases are driven by distorted beliefs; if instead they are caused by identity-preferences, companies could adapt to these consumer preferences by changing product attributes accordingly, e.g., choose more

"American" designs as a non-US car company. Our model could explain how patient treatment choices might be swayed by identity, emphasizing the need for tailored educational initiatives or treatment plans (Chan, 2022). These examples illustrate our model's broad applicability and relevance in understanding and influencing behavior through the lens of social identity.

This article presents three contributions to the economics of identity. Firstly, we bridge two separate strands of research: one focusing on the impact of identity on beliefs and the other on preferences, often disregarding or making assumptions about beliefs.[4] Conversely, another research stream has demonstrated that identity preferences—ignoring beliefs, controlling for, or assuming rational expectations—play a critical role in decision-making, as seen in the behavior of sports fans, investment choices, occupational choice and food consumption (Paul and Weinbach, 2009; Morewedge et al., 2018; Atkin et al., 2021; Oh, 2023). Our work bridges these strands by integrating the effects of identity-distorted beliefs (not assuming rational expectations) and preferences, offering a comprehensive perspective on how identity influences decision-making. Our findings demonstrate that belief distortions and preference-driven behaviors are each pivotal in decision processes. We extend beyond previous research by quantifying the contributions of beliefs and preferences on identity-relevant decisions, a previously unexplored area.

Secondly, we innovate methodologically by addressing the limitations of previous studies on identity. Several studies on identity rely on lab experiments, surveys, and ethnographic analyses (for a review of methods, see Abdelal, 2009), which provide valuable insights but often struggle to identify the causal impact of social identity and its influence on choice behavior. Our approach circumvents most of these concerns using large-scale lab-in-field experiments in diverse settings (Kenya and the UK). We employ both survey approaches of measuring identity and use formal structural modeling techniques combined with a revealed preference approach to infer how identity affects economic decision-making. Our theoretical and empirical approach is valuable for the literature as it provides a framework for understanding the implications of identity and demonstrates practical ways of quantifying identity-related biases in real-world scenarios. As such, our paper adds to a growing number of papers that structurally estimate models of social concerns and influences (DellaVigna et al., 2012, 2016; Karing, 2018; Donkor, 2021; Butera et al., 2022).

[4] Prior research has shown that group membership impacts beliefs and choices (Tajfel, 1970; Hogg, 1992; Chen and Li, 2009; Goette et al., 2006, 2012; Dimant, 2023) but also distorts beliefs about their environment (Cacault and Grieder, 2019; Bicchieri et al., 2023). For example, it is documented that partisan identity affects people's belief formation (Bauer et al., 2023) and their views on climate change, vaccines, and 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., 2023).

Thirdly, our research enriches the understanding of consumer choice by integrating social identity into economic decision-making. Like Atkin et al. (2021), we show that incorporating social identity into our analysis of consumer choice can enrich our understanding of how identity concerns influence economic behavior. For example, while traditional economics predicts that a price increase will lead to a fall in demand, for products closely tied to social identity, demand might be less elastic when consumers are willing to pay to maintain their social status (Heffetz, 2011; Charles et al., 2009; Bursztyn et al., 2020). This observation can explain why luxury goods sometimes see stable or even increased demand despite rising prices. The concept of social identity suggests that the perception of goods or assets as substitutes or complements can shift depending on how they fit within the narrative of a consumer's identity.[5] Indeed, our findings suggest that benefits associated with particular identities tend to generate complementarities.

The rest of the paper is structured as follows. After introducing the portfolio allocation model under identity concerns in Section 2, we delve into the specifics of our field experiments on soccer betting in the context of the English Premier League in Section 3. In Section 4, we present our reduced-form empirical findings from the experiments. Section 5 employs a semi-structural approach to assess identity-driven differences in decision-making after accounting for risk aversion, subjective beliefs, and differences in expected returns. Section 6 involves the structural estimation of the portfolio allocation model, and conducts counterfactual estimates, and heterogeneity analyses. In Section 7, we discuss the applications and implications of our framework and findings. Section 8 concludes.

2 A Model of Portfolio Allocation with Identity Concerns

We model an investment decision problem that accounts for agents' affiliations and identification with specific assets within their investment portfolio. In the model, agents exhibit affinity towards certain assets which are reflected in their investment decisions. Such affiliations may manifest in various forms: preferential stock holdings from one's employer, investing in one's own venture, or patriotic investments in domestic firms. A more nuanced understanding might include the propensity to invest in green energy companies due to environmental concerns or an inclination towards companies or institutions that support specific social causes aligned with an investor's personal values or ideologies.

The model postulates that identity-related concerns pertaining to an asset will lead to

[5]For example, someone with a strong pro-environmental identity might substitute stocks of a domestic company that is environmentally friendly with stocks of an environmentally friendly, foreign company. A patriotic investor, on the other hand, might instead substitute stocks of the same domestic and environmentally friendly company with stocks of another domestic company that is not environmentally friendly.

systematic differences in investment behavior compared to investments in assets devoid of such identity associations. For concreteness, we focus the model on how soccer fans bet on the potential outcomes of Premier League (PL) matches. This context ties directly to our field experiments, making it more straightforward to interpret our empirical findings. However, the model is broadly applicable.

We assume that fans seek to maximize their subjective expected utility when betting. Thus, when two teams participate in a match (home team versus away team), fan i faces the task of determining how to allocate a given budget across the potential outcomes of the game: h (home win), a (away win), or d (draw). Let $p_{i,j}$ represent fan i's subjective belief concerning the probability of outcome j occurring. Right before betting on a future match, fan i is presented with decimal odds denoted as θ_j (representing the number of wins per every dollar wagered on outcome j). Precisely, if fan i were to bet $x_{i,j}$ on outcome j at odds θ_j, her potential return is $x_{i,j} \times \theta_j$. We express the utility derived from the return on asset j as $u(\theta_j x_{i,j})$, where u represents a function that is strictly increasing and concave ($u' > 0$ and $u'' < 0$), thus reflecting the typical risk aversion behavior of investors.

In instances where fan i maintains neutrality towards the teams participating in a match, we express her betting decision problem as

$$\text{Max}_{x_{i,h},x_{i,a},x_{i,d}} U_i^N \quad = \quad p_{i,h}u(\theta_h x_{i,h}) + p_{i,a}u(\theta_a x_{i,a}) + p_{i,d}u(\theta_d x_{i,d}). \tag{1}$$

Conversely, when fan i supports the home team, her affiliation may introduce (i) distortions into her beliefs $(p_{i,h}, p_{i,a}, p_{i,d})$, and (ii) lead her to discount the payoffs associated with identity-incongruent outcomes (such as a draw or an away team win).[6] In this context, we represent her decision problem as follows:

$$\text{Max}_{x_{i,h},x_{i,a},x_{i,d}} U_i^S \quad = \quad p_{i,h}u(\theta_h x_{i,h}) + (1-\alpha_a)p_{i,a}u(\theta_a x_{i,a}) + (1-\alpha_d)p_{i,d}u(\theta_d x_{i,d}), \tag{2}$$

where $0 \leq \alpha_a, \alpha_d \leq 1$. This formulation implies that she assigns less weight to the utility derived from the potential returns in cases where the home team does not emerge victorious, i.e., either draws or loses to the away team.

The underlying psychological motivation in the presence of identity concerns lies in anticipating a distaste for gains stemming from identity-incongruent assets. This psychic distress translates into underweighting the expected utility associated with identity-incongruous sources. Given the increasing and concave nature of u, the underweighting of expected utility

[6]For clarity of exposition, we assume that fan i supports the home team or maintains a neutral stance concerning the teams competing in a given match. When she supports the home team, she always prefers a home team win over a draw or an away team win.

from identity-incongruent outcomes results in the following proposition:

Proposition 1. *If $0 \leq \alpha_a, \alpha_d \leq 1$, and a fan is a supporter of the home team, then her optimal portfolio allocation is such that $x^S_{i,h} \geq x^N_{i,h}$.*

This proposition posits that fan i will systematically overinvest in identity-congruent assets when her identity concerns lead her to undervalue identity-incongruent assets compared to a neutral stance. The proof of the proposition is presented in Appendix A. It is important to note that fan i is considered neutral when $\alpha_a = \alpha_d = 0$. The empirical validation of Proposition 1 constitutes the core objective of the structural and experimental analysis in this paper.

2.1 Portfolio Allocation

In standard expected utility theory, rational actors who are risk-neutral or risk-loving allocate their entire budget to the outcome with the largest expected reward, often resulting in corner solutions. Risk-averse agents in contrast would choose interior solutions, allocating positive amounts to all outcomes. In making these decisions, risk-averse agents consider the different returns on assets (governed by odds), their risk attitudes, and their subjective probabilities for different match outcomes. To capture these intricacies, we assume that individuals have CARA utility, denoted by

$$u(\theta_j x_{i,j}) = \frac{1 - e^{-r_i \theta_j x_{i,j}}}{r_i}, \text{ with } r_i \neq 0.$$

Here, r_i signifies the CARA risk aversion parameter of individual i.

2.1.1 Optimal Portfolio Allocation for Neutrals

The optimal budget allocation for a neutral fan, as determined by the first-order conditions (FOCs) from equation (1), is as follows:

$$x^N_{i,h} = \frac{\theta_d \theta_a}{\Theta} + \frac{1}{r_i} \underbrace{\left[\frac{(\theta_d + \theta_a)}{\Theta} \text{Ln}\left(\frac{p_{i,h}\theta_h}{p_{i,d}\theta_d}\right) - \frac{\theta_d}{\Theta} \text{Ln}\left(\frac{p_{i,a}\theta_a}{p_{i,d}\theta_d}\right) \right]}_{\Delta P_{i,h}} \tag{3}$$

$$x^N_{i,a} = \frac{\theta_d \theta_h}{\Theta} + \frac{1}{r_i} \underbrace{\left[\frac{(\theta_d + \theta_h)}{\Theta} \text{Ln}\left(\frac{p_{i,a}\theta_a}{p_{i,d}\theta_d}\right) - \frac{\theta_d}{\Theta} \text{Ln}\left(\frac{p_{i,h}\theta_h}{p_{i,d}\theta_d}\right) \right]}_{\Delta P_{i,a}} \tag{4}$$

Here, $\Theta = \theta_a \theta_h + \theta_d \theta_h + \theta_d \theta_a$ and the investment budget is normalized to one.

The first-order conditions provide an intuitive mechanism for budget allocation. Focusing on equation (3), the proportion of the budget wagered on a home victory hinges on the relative marginal returns of all possible match outcomes (the $\frac{\theta_d \theta_a}{\Theta}$ term) plus the inverse of the individual's CARA risk preference parameter multiplied by the subjective expected marginal returns of betting on the home versus betting on the away team—all modulated through $\Delta P_{i,h}$.

$\Delta P_{i,h}$ is the odds-adjusted difference in the logarithm of the subjective expected marginal returns between a home win and an away win, standardized by the expected marginal return of a draw. This mechanism ensures that a fan will allocate a more substantial portion of her budget to a home win if the subjective expected marginal return on this outcome is expected to be higher than that for an away win. Furthermore, the sensitivity of allocation to changes in $\Delta P_{i,h}$ is inversely proportional to the investor's risk aversion—the lower the risk aversion (smaller r_i), the greater the responsiveness.

An increase in the odds (θ_h) for a home win, for example, presents a dual effect: it directly amplifies the return on investments for a home win; however, it simultaneously weakens the incentive to allocate funds to a home win. This observation is due to an inherent insurance motive, which dictates that an increase in the expected utility of favorable outcomes (declining due to diminishing marginal utility) leads rational agents to redistribute their investments towards outcomes with higher marginal utility, reflecting a risk mitigation strategy.

2.1.2 Optimal Portfolio Allocation with Identity Concerns

We now turn to portfolio allocation decisions under identity concerns. Risk-averse agents influenced by identity concerns do not only assess financial returns and risk but also consider the alignment of their investments with their identities.

For analytical tractability, we assume $\alpha_d = \alpha_a = \alpha$, a simplification that does not compromise generality. This assumption implies that supporters of the home team weigh the utility derived from a draw or a home team loss uniformly when these outcomes do not match their preferred result—a home team win.[7]

Under identity concerns, fan i's optimal budget allocations derived via the FOCs from

[7] The primary objective of this study is to explore the broader implications and applications of our model of investment or consumption behavior under identity concerns rather than focusing on context-specific details of soccer like the distinction between a draw and an away win.

equation (2) are in line with Proposition 1:

$$x^S_{i,h} = x^N_{i,h} + \underbrace{\mathrm{Ln}\,(\alpha)\, r_i^{-1} \times \frac{-(\theta_d+\theta_a)}{\Theta}}_{\text{Overinvestment}} \tag{5}$$

$$x^S_{i,a} = x^N_{i,a} + \underbrace{\mathrm{Ln}\,(\alpha)\, r_i^{-1} \times \frac{\theta_d}{\Theta}}_{\text{Underinvestment}}. \tag{6}$$

That is, individuals tend to overinvest in identity-congruent assets when their identity concerns lead them to undervalue identity-incongruent assets compared to a neutral stance. To empirically test Proposition 1, will require an investment scenario meeting the following criteria: a) a measure or indicator of a person's identity for each investment decision, b) exogenous variation in the the tradeoff between consumption and identity utility, c) measures of individuals' subjective likelihoods and d) an objective benchmark of investing without identity concerns. We fulfill these prerequisites through two large-scale field experiments conducted with soccer supporters in the UK and Kenya. We present details of the experiments in the subsequent section.

3 Context of Soccer Betting and Experimental Design

Soccer betting is an optimal setting for our research, offering measurable indicators of team identity and clear trade-offs between consumption utility and identity concerns. Extant evidence shows that soccer fans are reluctant to bet against their favored team (Morewedge et al., 2018; Kossuth et al., 2020) and overconfident in predicting match outcomes (Chegere et al., 2022).[8] The setting allows us to change the relative price of the tradeoff between identity and consumption utility by experimentally varying the betting odds. Third, in the context of soccer betting, individuals make investment decisions that only differ as to whether they involve identity concerns or not. Apart from this, the decision-making environment remains constant in all other regards. We can thus cleanly analyze how identity concerns distort fans' investment decisions. Fourth, in the context of soccer betting, it is natural to elicit subjective beliefs for match outcomes and bookmaker's beliefs (as given by their published odds) serve as an objective benchmark for the likelihood of different match outcomes. Eliciting individuals' subjective beliefs about win, draw, and loss probabilities allows us to compare these to an objective benchmark and to use them in semi-structural and structural estimations.

[8] In previous work, betting behavior has been used to study, for example, corruption and savings behavior (Wolfers, 2006; Deutscher et al., 2017; Herskowitz, 2021).

3.1 Context: Betting Market

Beyond its research relevance, soccer betting is an economically significant activity, with its market size and growth rates highlighting the importance of understanding betting behaviors for market regulation. In 2021, sports betting was valued at $70.23 billion (about 30-40% of total gambling), with a predicted compound annual growth rate (CAGR) until 2028 of around 11%. Soccer betting accounts for more than 30% of revenues ($20.1 bn.) and also has the highest expected CAGR. In 2020, the sports betting market had ca. 197,500 employees which would translate to around 59,000 employees (30%) working in soccer betting.[9] The liberalization of betting markets and their continued growth therefore also raise the importance of understanding people's betting behavior and its implications for the optimal regulation of betting markets.

3.2 Recruitment and Participants' Decision Tasks

We recruited a sample of 2,410 soccer fans from Kenya and the UK before the beginning of the Premier League Season 2021/22. In Kenya, we recruited through Ajua, a marketing firm that regularly surveys young Kenyans for business customers. UK participants were recruited via Prolific. In both cases, participants were screened by stating that they were soccer fans and supported a Premier League Club. To get more details on their fan loyalty, we asked all participants before the season about their favorite team. During the PL season, we invited our pool of participants a few days before matchdays to place bets on three or four upcoming games.

We incentivized participants by paying them a base amount of $0.5 in Kenya and $0.6 in the UK for each matchday. Additionally, we offered a bonus based on participants' wagers, the odds, and the actual match outcomes. At the end of each survey round, we randomly selected one match per participant to determine their bonus. We informed participants after each matchday survey which match would affect their bonus. We gave each respondent 100 tokens to bet on each match. Under fair odds, participants could expect an average bonus of 100 tokens per matchday, amounting to KSh100 ($1.0) in Kenya and £2.0 (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 occurred.[10] We disbursed the earnings from the previous matchday a few days after each matchday.

Most participants completed our survey in around five minutes, which included two key questions about each match. Initially, we asked them to assign subjective likelihoods, ranging

[9] Read Polaris (2022) for more details.

[10] The maximum potential win was 3,770 KSh ($37.7) in Kenya and £75.4 in the UK.

from 1% to 98% and totaling 100%, for each of the three match outcomes. These likelihoods correspond to p_{ij} in our model. We chose not to incentivize belief elicitation to avoid potential biases like a stronger tendency towards 50%, which complex incentivization mechanisms might introduce, as Danz et al. (2022) demonstrates. We address potential measurement error in beliefs using an IV approach in our empirical analysis.

After eliciting subjective beliefs, 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 the 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 30 on "Win Man City", the interface would show 86 tokens under "Winnings" for "Win Chelsea", 132 for "Draw", and 104 for "Win Man City". This approach simplified the task for participants by eliminating the need for them to calculate their potential earnings.

3.3 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).

Fan support 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 6 teams cover 99% of supporters in Kenya and 79% of supporters in the UK. An exact breakdown of supporters by teams is shown in Appendix table A.1.

Team support is stable over time, as most participants have been supporting their favorite team for at least 10 years (as in Kenya) or for 20 years or longer in the UK.[11] While the reasons participants have come to support a specific 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, most participants and fans typically stick with it. This is also reflected in the ages of 15 (Kenya) and 11 (UK) at which the median participant started supporting their favorite team. This is 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

[11] 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.

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 Hotspurs.

3.4 Inferring Bookmaker Beliefs

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 the fair odds $1/p_j^{BM}$, then 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 whenever we infer bookmaker beliefs from a given set of odds.

3.5 Experimental Design

Our experimental design varied the relative returns for betting on the home team's loss versus its win. To do this in a standardized way across matches, for each match we calculated 7 sets of home and away odds θ_h^z, θ_a^z that would yield wedges $z \in \{-2.5, -1.5, -.5, 0, +.5, +1.5, +2.5\}$,[12] where

$$z = p_a^0 * \theta_a^z - p_h^0 * \theta_h^z \tag{7}$$

Ahead of every matchday, we applied the formula (described in detail in Appendix section B.3) for the respective wedges z to the benchmark odds that we obtained from the German betting company "Tipico". We randomized odds at the individual-match-level, such that for each match, respondents would be shown randomized sets of odds that corresponded to one of the 7 wedges. Each set of odds would be shown with an equal probability of $\frac{1}{7}$. In Table A.3, we show the randomized odds (and the underlying change in assessed likelihood) for different sets of benchmark odds.

In addition to variation in fan identity, we thus have two layers of exogenous variation in relative returns of betting on a win versus a loss: within-person variation in returns for the

[12] 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.

same team across matches and across-person variation within matches.

4 Reduced-Form Analysis

First, we investigate the effect of identity on wagers and on beliefs regarding the probability of match outcomes. Analysis of our experimental data reveals a significant effect of team allegiance: fans increase their budget allocation to bets involving their favored teams by 22% compared to bets on neutral matches. Moreover, fans' beliefs about their team's chances of winning a match are 10% to 18% higher than those of their neutral counterparts.

4.1 Portfolio Budget Allocation: Identity-Specific Betting

We begin our analysis by exploring the impact of fan identity on investment decisions. Figure 2 displays a summary of the betting decisions of soccer fans from the UK and Kenya. Betting patterns among fans from the UK and Kenya show remarkable similarities: supporters of the home team allocate 10 percentage points more, while away team supporters allocate 8 percentage points less, to a home win compared to neutral bettors.

To delve deeper into the discrepancies in betting behavior based on fan identity, we adjust for individual and match-related differences by incorporating individual and match-fixed effects in our analysis. Table 2 presents these conditional betting outcomes. Supporters of the home (away) team allocate 8.5 (six) percentage points more (less) of their budget to a home win compared to neutrals, in both the UK and Kenya. For an away team win, away (home) team supporters in the UK and Kenya allocate about eight (six) percentage points more (less) of their budget than neutrals.

When fans support the home or away team, there is a notable difference in their betting behavior compared to neutral fans.

4.2 Identity-Distorted Beliefs

Next, we examine the interplay between team allegiance and beliefs regarding the outcomes of PL matches. Figure 3 illustrates the CDF of participants' predictions about match outcomes, differentiating between team allegiances and neutrals. Supporters consistently exhibit greater optimism about the likelihood of their team's victory prospects than neutrals, especially when juxtaposed against rival team fans. These patterns persist even after adjusting for individual and match fixed-effects (as presented in Table 3).

Comparing neutral fans' predictions with bookmakers' odds, we observe alignment in the Kenyan sample, where neutral fans estimate the likelihood of a home team winning at 39%, mirroring the bookmakers' prediction. However, their neutral UK counterparts anticipate a marginally lower chance of a home victory at 39% than the bookmaker's 41%. The average prediction of the likelihood of an away team's victory stands roughly at 39 to 40% for neutrals from both countries, as opposed to the bookmaker at 36%.

Supporters of both home and away teams in the UK (and Kenya) report higher probabilities for their team winning a match by margins of 4.4 and 4.0 (7.0 and 5.2) percentage points, respectively, compared to neutral fans. Supporters' forecasts of their team emerging victorious in a game is between 10% and 18% above the projections of neutral fans.

In conclusion, our analysis indicates that soccer fans in both the UK and Kenya display an optimistic bias in estimating their supported team's chances of winning. This bias is evident when compared to the neutral perspectives of non-supporting spectators and bookmakers' forecasts. On the other hand, supporters also underestimate the chance of their team losing, both at home and away. Compared to bookmakers' forecasts, especially away fans are overoptimistic about their team's chances of winning.

5 Semi-Structural Analysis of Differences in Portfolio Allocation

According to neo-classical economics, rational investors' portfolio decisions are governed solely by subjective expectations of asset returns and individual risk preferences. After conditioning on these factors, there should be no systematic differences in investors' budget allocation across portfolio assets. This section empirically evaluates whether fan identity sways investment decisions beyond subjective beliefs, potential returns (betting odds), and risk preferences. We propose the following null hypothesis:

Hypothesis 1. *Team allegiance has no impact on investment decisions compared to neutrals after accounting for subjective beliefs, betting odds, and risk preferences.*

Testing this hypothesis addresses whether heterogeneity in subjective beliefs or individual risk preferences is sufficient to explain the neutral-supporter gap we observe in bets. We leverage the FOCs for neutral bettors (equations (3) and (4)) as the baseline portfolio allocation when fan identity is inconsequential. We then assess whether the empirical structural components gleaned from these equations exhibit patterns that correlate with fan loyalty.

Equations (3) and (4) show that budget allocations hinge on betting odds, personal beliefs, and risk preferences. We directly elicit subjective beliefs and exogenously vary betting odds in our experiment. Thus, the CARA risk preference parameter, r_i, remains the component for estimation.

5.1 Estimating Individual Risk-Preference

For each match, denoted by m, both equations (3) and (4) jointly identify r_i, rendering it over-identified. Using the panel data of betting decisions for individual i, we estimate equations (3) and (4) simultaneously in a linear regression to identify the CARA risk preference parameter r_i. The estimation equation is:

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

The outcome variable is individual i's budget share $x_{j,m}$ allocated towards outcome j (home or away win), in match m, minus the empirical analogue of the first component of the theoretical portfolio allocations (FOCs for neutral bettors), the ratio of exogenous individual-match specific betting odds. The independent variable is $\Delta P_{j,m}$, and $\varepsilon_{j,m}$ is the residual term. Thus, variation in individual i's subjective beliefs and exogenous betting odds across matches (m) and outcomes (j) identifies $\beta = \frac{1}{r_i}$ via the term $\Delta P_{j,m}$.

There is a concern that misreporting subjective beliefs could lead to inaccurate measurement of $\Delta P_{i,j,m}$, biasing our estimate of β. We address two potential measurement errors. The first is random errors in reports of subjective beliefs. In this scenario, the random nature of the error could introduce attenuation bias in the estimate of $(\frac{1}{r_i})$, leading to an overestimation of risk preferences. We can tackle i.i.d measurement error by using an instrumental variables approach. The second potential measurement error is if supporters exaggerate their team's chances of winning to the researcher (e.g., because of self-image concerns) but use a more accurate or less skewed belief when deciding on portfolio allocation.The direction of bias due to systematic misreporting is not straightforward.[13] Thus, we use the subset of matches when individuals are neutral and an instrumental variables strategy to estimate r_i. This strategy addresses random measurement error and sidesteps potential systematic misreporting. In the IV first stage for estimating r_i, we construct $\Delta P^{BM}_{j,m}$ using bookmaker beliefs and then use it as an instrument for the potentially mismeasured $\Delta P_{i,j,m}$.

5.2 Empirical Test for Null Hypothesis

Having computed all structural components of the FOCs, we investigate whether fan allegiance systematically influences portfolio allocations. We estimate the following regression

[13] For example, an increase in $p_{i,h,m}$ implies a decrease in $p_{i,a,m}$ and/or $p_{i,d,m}$, affecting the log terms in the numerator of $\Delta P_{i,j,m}$. Thus, the overall impact on r_i depends on how the changes in the numerator affect the whole equation. Focusing on equation (3), if a respondent inflates $p_{i,h}$, the numerator increases more relative to r_i, and the estimate of $\frac{1}{r_i}$ would be biased towards zero, hence an overestimation of r_i. The opposite is true if $p_{i,a,m}$ is exaggerated. Therefore, addressing the systematic misreporting, especially when fan loyalty is a factor, is trickier.

equation:

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

The outcome variable is individual i's budget share $x_{i,j,m}$ allocated towards outcome j in match m, minus the empirical analogue of the components of the theoretical portfolio allocation decision (the right hand side of the FOCs of neutral bettors in equations (3) and (4)). $\hat{r}_i$ is the CARA risk aversion parameter for individual i estimated from equation (8) above. The coefficients γ_i and γ_2 on the supporter indicators should be statistically indistinguishable from zero if our null hypothesis holds, thereby confirming that team allegiance does not affect investment choices. We control for individual and match fixed effects, denoted by I_i and λ_m, respectively.

5.3 Semi-Structural Results

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

Figure 4 presents a visual representation of the empirical components of the theoretical portfolio budget allocation towards a home win (equation (3)) for bettors in the UK (top panel) and Kenya (bottom panel). The figure is a binned scatter plot detailing how bets on the home win vary with the level of $\Delta\hat{P}_{i,j}$ (fitted values of $\Delta P_{i,j}$), separately for home team supporters (blue), away team supporters (red), and neutrals (yellow). We include fitted lines by fan identity, the slopes of which are estimates of the inverse of identity-specific CARA risk preference parameters.

Under the null hypothesis, the plotted points should overlap with no systematic differences by team allegiances. The results for the UK and Kenya in Figure 4 indicate the contrary. First, the points representing budget shares allocated towards a home win by home team supporters are systematically higher than neutrals' shares, whose allocated shares in turn are higher than those of away team supporters. This observation suggests that home team supporters consistently bet more on home win than neutral fans, whereas away supporters do the opposite.

Second, the slopes of all the fitted lines by fan identity are positive, suggesting that soccer fans are, on average, risk-averse. Notably, while there are wedges between the fitted lines by fan identity, the lines are more or less parallel, which suggests that risk aversion is similar across fan identity and, therefore, likely not the explanation for the systematic differences in betting behavior between home, neutral, and away team supporters.

For a more formal evidence, we estimate an analogous regression to Figure 4, where we assume homogeneous risk preferences, and instrument $\Delta P_{i,j,m}$ using $\Delta P^{BM}_{j,m}$ in a TSLS regression that accounts for individual and match fixed effects.[14] The results are shown in Table 4, where we test for identity-specific differences in betting behavior by using indicator variables of team allegiance. The first (third) column represents the findings illustrated in Figure 4's top (bottom) panel for the UK (Kenya). The second and fourth columns report the findings for differences in portfolio budget allocations for an away-win for the UK and Kenyan samples, respectively.

The results show that the differences we observe between home and away team supporters' bets compared to neutrals are statistically significant. The estimates also yield a homogeneous CARA risk preference r value close to 0.01, suggesting that individuals in both samples are risk-averse on average.

We now turn to estimating and accounting for individual risk preferences in our test for the null hypothesis. We estimate r_i via equation (8), as discussed in section 5.1 above. Figure A.3 presents the distribution of r_i estimates for our samples from the UK and Kenya. We observe a higher average risk aversion parameter $r_i = 0.021$ in the Kenyan sample compared to $r_i = 0.015$ in the UK sample.

We estimate equation (9), which tests our null hypothesis and accounts for individual-specific risk preferences. We report the findings in Table 5. We find that the differences in portfolio budget allocation based on fan identity amplify, roughly doubling when we account for individual-specific risk preferences.

As a robustness check, we replace (impute) respondents' beliefs with bookmaker's beliefs and re-estimate the specifications in Table 6. 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.

[14]The corresponding figures for budget allocation towards away win are shown in Appendix Figure A.5.

6 Structural Analysis of Differences in Portfolio Allocation

This section measures how supporters' and neutrals' portfolio choices differ, factoring in more than just their subjective beliefs and risk preferences. The portfolio allocation model with identity concerns indicates that individuals undervalue (or underweight) the expected utility from gains from their supported team not winning a match. We operationalize individuals' undervaluation of potential returns from sources inconsistent with their identity compared to neutral assets as an ad-valorem tax α in the model.

Introducing identity concerns in investment decisions in the model produces an extra set of FOCs for supporters (equations (5) and (6)). Proposition 1 states the implications of these FOCs, asserting that supporters excessively invest in their team's victory compared to neutrals because of the identity incongruence tax (identity tax for short).

We empirically estimate the identity tax by employing equations (5) and (6); We set α to zero for bets on neutral games. We follow the same estimation procedure as in the semi-structural section above by addressing measurement error in $\Delta P_{i,j}$ using an IV approach and estimating r_i on individuals' neutral matches.

We find that supporters markedly devalue the expected utility from payoffs resulting from their team not winning a match. We report the structurally estimated identity tax rates in Table 7, differentiating between the UK and Kenya. Assuming uniform risk preferences, we estimate an identity tax of approximately 11.8% for the UK and 9.4% for Kenya (first two columns of Table 7). The UK and Kenyan participants exhibit comparable risk preferences r, roughly at 0.01.

When we allow for individualized risk preferences (columns three and four of Table 7), our estimates show the identity tax increases to 16.6% in the UK (from 11.8%) and 27.2% in Kenya (from 9.4%). These tax rates correspond to utility weights of 0.83 for the UK and 0.73 for Kenya for the expected earnings from identity-incongruent outcomes. An explanation for the higher estimates of the identity tax under heterogeneous risk preferences is that it accounts for above-average risk aversion among a significant share of the sample (especially in Kenya, see Figure A.3). Because deviating by a fixed amount from the neutral benchmark is particularly costly in utility terms for individuals with higher risk aversion, such deviations are rationalized by a higher identity tax.

6.1 Decomposing Identity Investment Differentials

To assess the relative importance of identity-induced differences in beliefs and identity-based preferences for investment decisions, we use equations (3) and (5) to generate various counterfactuals using individual risk aversion estimates. As the equation makes clear, investment x^N depends on risk aversion r_i, beliefs p, and odds θ. In our decomposition exercise, we set $\theta = (3, 3, 3)$ for all counterfactuals.[15]

We first construct the counterfactuals to assess the difference in bets due to differences between neutral beliefs p^N and supporters' beliefs p^S. Following equation (3) and using the supporter's estimated risk aversion, we want to calculate what supporters (with α set to zero) would bet with identity-specific beliefs and with neutral beliefs. Since beliefs enter x^N non-linearly, we cannot simply replace the beliefs of supporters with the mean belief of neutrals to generate the counterfactual. To assign neutral (identity-specific) beliefs, we instead take the set of supporters for each match and sample (with replacement) from the distribution of beliefs about a home win held by neutrals for the corresponding match. We then calculate the average investment under neutral (identity-specific) beliefs for each match and average these match-specific averages over all matches. $\bar{x}^N(p^S, r) - \bar{x}^N(p^N, r)$ then corresponds to the impact of identity on investment through the belief channel.

To construct the counterfactual for the impact of preferences through the identity tax α in our model, we use equation (5) to calculate $\bar{x}^S(p^S, r)$. This decomposition does not depend on the order (beliefs vs. preferences), since beliefs do not enter the term involving the identity tax in equation (5).

Table 8 displays the results for this decomposition. The upper panel shows the results for the UK sample. We find that identity causes UK supporters to invest 24% more into their supported team's win. The decomposition shows that one-third of the effect comes from differences in beliefs between neutrals and supporters, while the other two-thirds are due to identity preferences. The bottom half displays the results for Kenya. While the overall magnitude in Kenya is about the same (a 21.5% difference in investments in the supported team's win), our estimates indicate that for Kenya, about half is due to differences in beliefs, whereas the other half is due to identity preferences.[16]

[15] These are the fair odds if all outcomes were equally likely and provide a useful benchmark. If $\theta = (3, 3, 3)$ and $p = (1/3, 1/3, 1/3)$, then individuals would invest $x_k^N = 100/3$ into each of the outcomes $k = h, d, a$.

[16] The absolute and relative contribution of identity preferences to differences in investment shares is thus lower in Kenya than in the UK. While higher levels of risk aversion in Kenya contributed to a higher estimate of α for Kenya, higher levels of risk aversion also reduce the impact of a given α on investment shares. More distorted beliefs among Kenyan supporters further reduce the relative contribution of identity preferences.

6.2 Heterogeneity in Identity Preferences

We examine the variations in identity-specific preferences among different teams and their fans, recognizing that supporters of lower-ranked teams likely exhibit distinct identity-specific preferences compared to those backing top-tier teams. We focus on quantifying the strength of the identity tax conditional on a team's rank in the previous season as well as conditional on the outcome of a team's most recent match.[17]

In our analysis, we parameterize $\alpha_{k,m} = \delta_k + \eta \cdot win_{k,m-1}$ in the FOCs (5) and (6). Here, δ_k represents a team-specific component in α, and $win_{k,m-1}$ is a binary variable denoting team k's victory in the previous $m-1$ match. Thus, η indicates how a team's short-run performance influences the identity incongruence tax. Notice that our structural estimation controls for differences in beliefs that might be associated with differences in long-term or short-term performance. Hence, the parametrization of α allows us to identify team-specific and short-run influences on the valuation of identity-incongruent outcomes.

We estimate the identity taxes specific to each of the top 17 teams from the 2020/21 season that were also present in the 2021/22 season, the season we are analyzing. We then explore the relationship between team ranking and identity tax. This analysis primarily considers UK fans, as the sample of supporters for each team in the league is large enough to allow us to estimate most team-specific identity taxes somewhat accurately.[18]

Table 9 displays our results, revealing significant heterogeneity in identity incongruence tax rates across teams. In particular, we uncover a strong pattern in the heterogeneity: teams at the lower end of the ranking exhibit a marked increase in the estimated identity tax, as Figure 5 illustrates. The highest-ranked team in our sample has an estimated identity tax of 12%, and the lowest-ranked team has an identity tax of 47%, marking almost a fourfold increase. These findings indicate that the identity tax is particularly strong when one's team is not doing well or an underdog, i.e., supporters of lower-ranked teams feel especially "bad" about benefiting from identity-incongruent outcomes. Psychologically, this could stem from stronger internal norms against benefiting from one's team's misfortunes when the team is struggling. The act of betting on a poorly performing team may also carry increased self-signaling value as it more clearly demonstrates loyalty – it is evident that the support through betting is not based on the team's success, but in spite of its poor performance.[19]

[17] The Premier League comprises 20 teams which are ranked by their number of points. The last three teams face relegation to the English Football League (EFL), with their slots taken by the EFL's top three teams in the same season.

[18] In Kenya, more than 90% of participants support one out of only four teams (Manchester United, Arsenal, Chelsea, and Manchester City), leaving an insufficient number of observations for most remaining teams.

[19] One could have also expected that supporters opportunistically choose to support their "high-performing" identity more strongly, but this is not what the average supporter appears to do.

Alternatively, the same pattern could come about if only the "strongest" supporters (as measured by their identity tax) stay loyal to their team if it is doing poorly, or a majority of the supporters of good teams are so-called glory-hunters with low identity taxes (weaker identity preferences). These alternative hypotheses find little support in our data. Table 1 confirms that our subjects have maintained allegiance to their teams for over 23 years on average, undermining the notion of 'glory-hunting' supporters with transient affiliations. Moreover, Table 9 shows that temporary fluctuations in a team's success also affect the identity incongruence tax in the same way: the point estimate of η is negative and significant. If a team won the previous game, the identity incongruence tax is 9.6% lower in the current game. Taken together, this evidence favors the interpretation that a team's misfortune strengthens the norm against profiting financially from one's team's adversity.

We provide more details on the team-level decomposition in Table A.2. There are several other changes besides an increase in the identity tax when comparing the results for low-ranked teams to those of high-ranked teams: the average stakes under neutral beliefs decline from around 50 to 60 to between 20 and 30, the absolute and relative difference in stakes due to identity (beliefs and preferences) increases, and the share of these differences due to identity-specific beliefs appears to decline.

7 Discussion

In this section, we discuss potential alternative interpretations and concerns about our results and illustrate how our theoretical model is applicable in various scenarios where identity plays a key role in economic decisions. In particular, our approach provides a methodology for distinguishing between identity-motivated and purely financially motivated investments. In addition, we elaborate on how our theoretical framework may be used in crafting business strategies or regulations when individuals have identity concerns.

7.1 Concerns and Alternative Interpretations

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). Our empirical outcomes do not corroborate the emotional hedging hypothesis as the dominant strategy. We estimate a positive α, suggest an overarching propensity for identity-consistent investments. This behavior aligns with the overinvestment propensities predicted by our model, specifically for identity-congruent assets (Proposition 1).

Arbitrage opportunities and small-stakes risk aversion. Our experiment can present arbitrage opportunities for survey participants. Suppose agents are devoid of identity concerns and are paying attention to or 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 finding risk aversion for smaller stakes is 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 generated. It is conceivable that beliefs arise as a function of identity-specific preferences (motivated beliefs). Should this interrelation hold, our decomposition offers a conservative estimation of identity preference-driven behavior. In essence, our findings may understate the degree to which identity concerns skew investment decisions, as hypothesized in our model.

7.2 Applications

Consumer Finance: Investment and Savings Decisions. The principles outlined in our model reveal significant implications for consumer finance, particularly concerning investment and savings decisions. Consumers' affinity for specific financial assets or products—akin to the loyalty displayed by soccer fans—may influence their choices in investment or savings vehicles. For example, an investor may exhibit disproportionate favoritism towards domestic equities or bonds for patriotic reasons. This proclivity may yield an 'identity-tax' on portfolio diversity. Should surveys or empirical analyses reveal that investor predispositions stem from misconceptions, educational campaigns presenting objective investment vehicle comparisons might mitigate biases born from misinformation. Conversely, if such biases are deep-seated preferences, policies may entail crafting diversified, government-mandated retirement portfolios that obviate the need for active asset selection, thus preserving identity while promoting diversification.

Policymakers should thus consider identity as an important driver of investment behavior when designing financial products and retirement plans. For instance, introducing diversified retirement savings plans that minimize active decision-making could cater to identity concerns while promoting financial stability and portfolio diversification.

Industrial Organizations and Marketing: Firm Competitive Strategy. Our model offers a nuanced perspective on the influence of identity-related concerns on investment choices that can be extrapolated to understand consumer decision-making in firm competitive strategies. Within the framework, consumer affiliations and identity affiliations—similar to soccer fans' betting behavior on favored teams—may manifest as brand loyalty or preference for products originating from an affiliation such as one's own country. These affiliations may produce systematic deviations in purchasing behavior, mirroring the identity-related distortions in our portfolio allocation model. For example, consider a U.S.-based consumer who systematically prefers American cars due to a strong nationalistic identity. Firms could measure the extent of this identity-tax and distorted beliefs through controlled experiments or consumer market research by offering a diversified set of choices that range from domestic to foreign brands.

Understanding whether this bias is due to distorted beliefs or genuine preference has crucial implications for firm strategy. If distorted beliefs are at play, as our model suggests can happen, the firm may opt for an "informative-advertising" strategy to correct misconceptions and neutralize identity distortions. On the other hand, if the bias is found to be a genuine preference, strategies such as price cuts or enhancements in product attributes targeted toward these identity-related preferences may be more effective. For instance, a foreign car manufacturer might incorporate design elements that resonate with American cultural symbols to appeal to the U.S. consumer's nationalistic identity—akin to recalibrating the α parameter (identity tax) in our model.

Occupational and Labor Market Mobility. Confronted with the decline of certain industries or regions, policymakers are often at a crossroads regarding facilitating workforce mobility. Consider the protracted downturn of coal mining in Germany's Ruhr region and the resultant occupational inertia. Our approach could elucidate the underlying causes of this inertia and inform policies to enhance societal welfare. If misperceptions about future earnings are a barrier, providing accurate information may stimulate mobility. Conversely, suppose a high identity-tax is the underlying deterrent. Policy responses may vary if the tax is a cognitive bias or a legitimate preference. If the former, policymakers might want to make transitioning relatively more appealing. If the latter, measures could be taken to mitigate identity loss, such as subsidizing employment in declining sectors until retirement or fostering new identity constructs within transitioning communities.

Healthcare: Patient Treatment Choices. Our model's insights extend to healthcare, where patient treatment choices can be influenced by identity concerns (Chan, 2022). Similar to fans' betting behaviors, patients may display biases toward treatments from providers or pharmaceuticals that align with their personal or cultural identity. A patient might favor a treatment developed or endorsed by their community. Healthcare providers and pharmaceutical companies could quantify the identity-tax associated with these biases through patient surveys or treatment outcome studies. If biases are due to misinformation, providers could engage in educational initiatives to clarify treatment efficacy and safety. If preferences are genuine, treatment plans could be tailored to enhance compliance. For example, a healthcare provider could emphasize the R&D behind a treatment to appeal to patients' community pride, adjusting the perceived identity-tax and improving treatment uptake.

In sum, our model not only offers a nuanced understanding of the role of identity in economic decision-making but also serves as a foundational tool for applying these insights across various sectors to enhance both individual and collective welfare. As future work continues to validate and refine this model through empirical research, the potential for tailored, identity-conscious policy interventions will become increasingly tangible.

8 Conclusions

This study introduces a framework that integrates social identity into economic decision-making, revealing a tendency for individuals to prefer identity-congruent consumption and assets. This integration enriches the subjective expected utility model with a deeper understanding of identity's role in shaping preferences.

Through lab-in-field experiments in Kenyan and UK soccer betting, our empirical analysis confirms that fans driven by identity-based beliefs and preferences allocate more funds to bets favoring their teams compared to when they are neutral. This demonstrates the tangible impact of identity on economic behavior. These findings persist even when accounting for individual risk preferences and subjective beliefs about asset returns, underscoring the influence of identity on economic choices.

Methodologically, we innovate by combining lab-in-field experiments with structural modeling techniques, allowing us to infer the impact of identity on economic decision-making more accurately. This approach addresses the limitations of previous studies, which often relied on laboratory experiments or ethnographic studies, providing a more robust and comprehensive understanding of identity's role in shaping economic behavior.

Our findings highlight the importance of incorporating identity considerations in policy design and research into the relationship between social identity and economic behavior across diverse contexts. Future investigations could explore the model's applicability to other decision-making scenarios and refine the methodology to capture more nuanced aspects of identity-influenced decision-making.

References

Abdelal, Rawi, *Measuring identity: A guide for social scientists*, Cambridge University Press, 2009.

Akerlof, George A and Rachel E Kranton, "Economics and identity," *The Quarterly Journal of Economics*, 2000, *115* (3), 715–753.

Allcott, Hunt, Levi Boxell, Jacob Conway, Matthew Gentzkow, Michael Thaler, and David Yang, "Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic," *Journal of Public Economics*, 2020, *191*, 104254.

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

Bauer, Kevin, Yan Chen, Florian Hett, and Michael Kosfeld, "Group Identity and Belief Formation: A Decomposition of Political Polarization," *Working Paper*, 2023.

Bicchieri, Cristina, Eugen Dimant, and Silvia Sonderegger, "It's not a lie if you believe the norm does not apply: Conditional norm-following and belief distortion," *Games and Economic Behavior*, 2023, *138*, 321–354.

Bohren, J Aislinn, Kareem Haggag, Alex Imas, and Devin G Pope, "Inaccurate statistical discrimination: An identification problem," *Review of Economics and Statistics*, 2023, pp. 1–45.

Bursztyn, Leonardo, Alessandra L. Gonzalez, and David Yanagizawa-Drott, "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia," *American Economic Review*, October 2020, *110* (10), 2997–3029.

Butera, Luigi, Robert Metcalfe, William Morrison, and Dmitry Taubinsky, "Measuring the welfare effects of shame and pride," *American Economic Review*, 2022, *112* (1), 122–68.

Butler, Jeffrey V, "Identity and the theory of the firm," in "Handbook of Behavioral Industrial Organization," Edward Elgar Publishing, 2018.

Cacault, Maria Paula and Manuel Grieder, "How group identification distorts beliefs," *Journal of Economic Behavior & Organization*, 2019, *164*, 63–76.

Cassan, Guilhem, Daniel Keniston, and Tatjana Kleineberg, "A division of laborers: Identity and efficiency in India," Technical Report, National Bureau of Economic Research 2021.

Chan, Alex, "Discrimination against Doctors: A Field Experiment," *Working Paper*, 2022.

Charles, Kerwin Kofi, Erik Hurst, and Nikolai Roussanov, "Conspicuous consumption and race," *The Quarterly Journal of Economics*, 2009, *124* (2), 425–467.

Chegere, Martin, Paolo Falco, Marco Nieddiu, Lorenzo Pandolfi, and Mattea Regina Stein, "It's a Sure Win! Experimental evidence on overconfidence in betting behavior," *CSEF WORKING PAPERS*, 2022, *655.*

Chen, Yan and Sherry Xin Li, "Group identity and social preferences," *American Economic Review*, 2009, *99* (1), 431–57.

Danz, David, Lise Vesterlund, and Alistair J Wilson, "Belief elicitation and behavioral incentive compatibility," *American Economic Review*, 2022, *112* (9), 2851–2883.

DellaVigna, Stefano, John A List, and Ulrike Malmendier, "Testing for altruism and social pressure in charitable giving," *The quarterly journal of economics*, 2012, *127* (1), 1–56.

_ , _ , _ , and Gautam Rao, "Voting to tell others," *The Review of Economic Studies*, 2016, *84* (1), 143–181.

Depetris-Chauvin, Emilio, Ruben Durante, and Filipe Campante, "Building nations through shared experiences: Evidence from African football," *American Economic Review*, 2020, *110* (5), 1572–1602.

Deutscher, Christian, Eugen Dimant, and Brad R Humphreys, "Match fixing and sports betting in football: Empirical evidence from the German Bundesliga," *Working Paper*, 2017.

Dimant, Eugen, "Hate Trumps Love: The Impact of Political Polarization on Social Preferences," *Management Science*, 2023.

_ , Fabio Galeotti, and Marie Claire Villeval, "Motivated Information Acquisition and Social Norm Formation," *Working Paper*, 2023.

Donkor, Kwabena, "The Economic Value of Norm Conformity and Menu Opt-Out Costs," *Available at SSRN 3955553*, 2021.

Egan, Patrick J and Megan Mullin, "Climate change: US public opinion," *Annual Review of Political Science*, 2017, *20*, 209–227.

Foad, Hisham S, "Equity home bias and the Euro," *The Quarterly Journal of Finance*, 2012, *2* (01), 1250004.

Gadarian, Shana Kushner, Sara Wallace Goodman, and Thomas B Pepinsky, "Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic," *Plos one*, 2021, *16* (4), e0249596.

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

_ , _ , _ , and Matthias Sutter, "Competition between organizational groups: Its impact on altruistic and antisocial motivations," *Management science*, 2012, *58* (5), 948–960.

Heffetz, Ori, "A test of conspicuous consumption: Visibility and income elasticities," *Review of Economics and Statistics*, 2011, *93* (4), 1101–1117.

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

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

Karing, Anne, "Social signaling and childhood immunization: A field experiment in Sierra Leone," *University of California, Berkeley Working Paper*, 2018.

Khaw, Mel Win, Ziang Li, and Michael Woodford, "Cognitive imprecision and small-stakes risk aversion," *The review of economic studies*, 2021, *88* (4), 1979–2013.

Kossuth, Lajos, Nattavudh Powdthavee, Donna Harris, and Nick Chater, "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*, 2020, *171*, 35–58.

Kurschilgen, Michael, "Moral awareness polarizes people's fairness judgments," *Social Choice and Welfare*, 2023, *61*, 339–364.

_ **and Isabel Marcin**, "Communication is more than information sharing: The role of status-relevant knowledge," *Games and Economic Behavior*, 2019, *113*, 651–672.

Manski, Charles F, "Measuring expectations," *Econometrica*, 2004, *72* (5), 1329–1376.

Meulbroek, Lisa, "Company stock in pension plans: How costly is it?," *The Journal of Law and Economics*, 2005, *48* (2), 443–474.

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

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

Park, Chan Su and Vern Srinivasan, "A survey-based method for measuring and understanding brand equity and its extendibility," *Journal of marketing research*, 1994, *31* (2), 271–288.

Paul, Rodney J and Andrew P Weinbach, "Sportsbook behavior in the NCAA football betting market: Tests of the traditional and Levitt models of sportsbook behavior," *The Journal of Prediction Markets*, 2009, *3* (2), 21–37.

Polaris, Market Research, "Sports Betting Market Share, Size, Trends, Industry Analysis Report," *Market Research Report*, 2022.

Rabin, Matthew, "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

Tajfel, Henri, "Experiments in intergroup discrimination," *Scientific american*, 1970, *223* (5), 96–103.

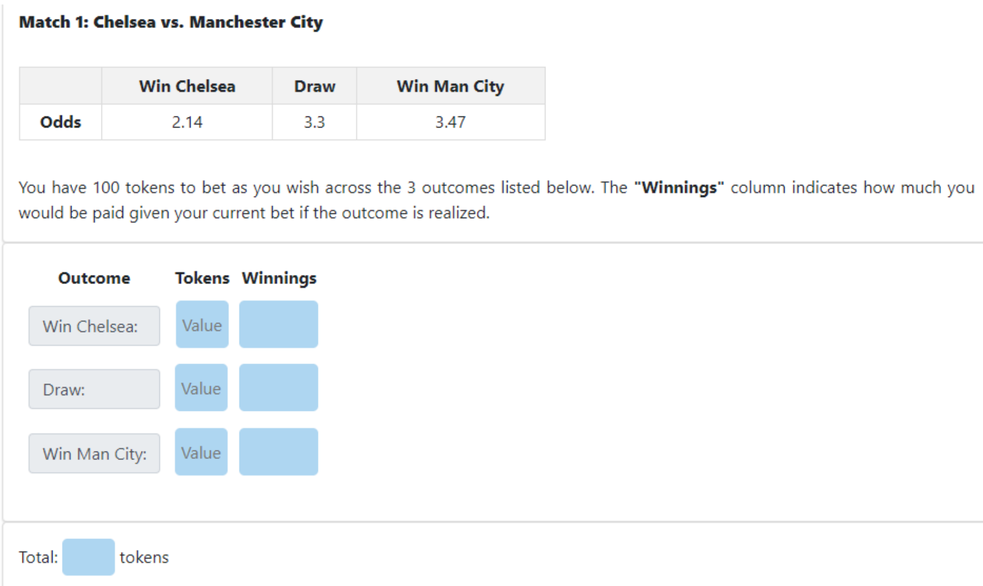
Wolfers, Justin, "Point shaving: Corruption in NCAA basketball," *American Economic Review*, 2006, *96* (2), 279–283.

Zerbib, Olivier David, "A sustainable capital asset pricing model (S-CAPM): Evidence from environmental integration and sin stock exclusion," *Review of Finance*, 2022, *26* (6), 1345–1388.

Figures and Tables

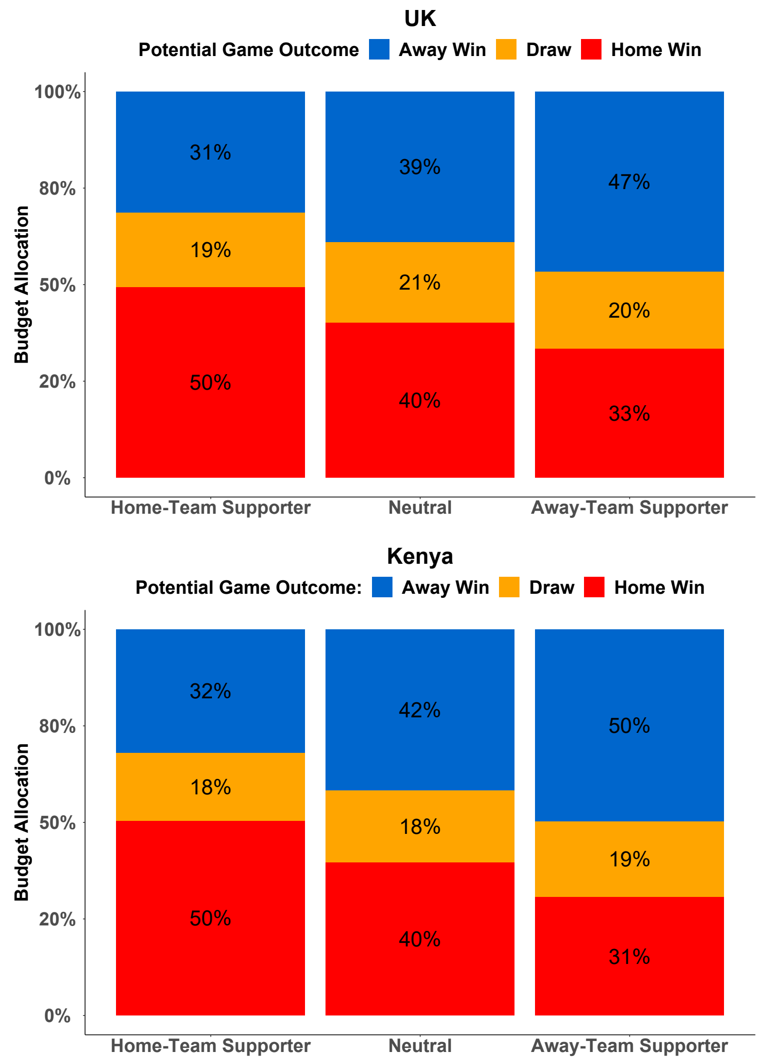
Figures

Figure 1: Interface for Participants Placing Bets



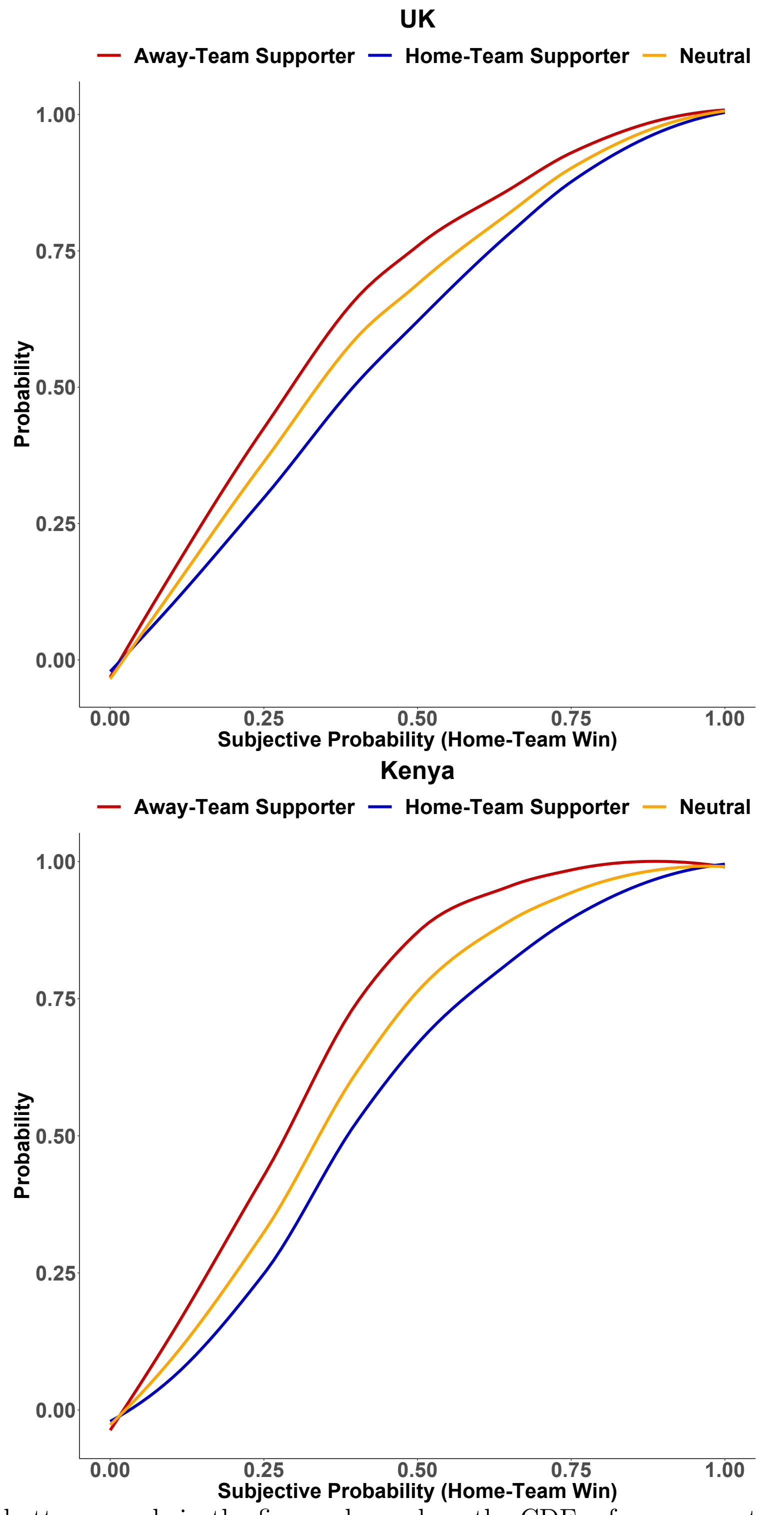
Notes: This figure displays the interface for participants placing bets. They are first informed about the odds for a home win, a draw, and an away win. Below, participants can then type in how many tokens they would like to place on the 3 outcomes. Next to their allocated number of tokens under "Winnings",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

Figure 2: Portfolio Allocations



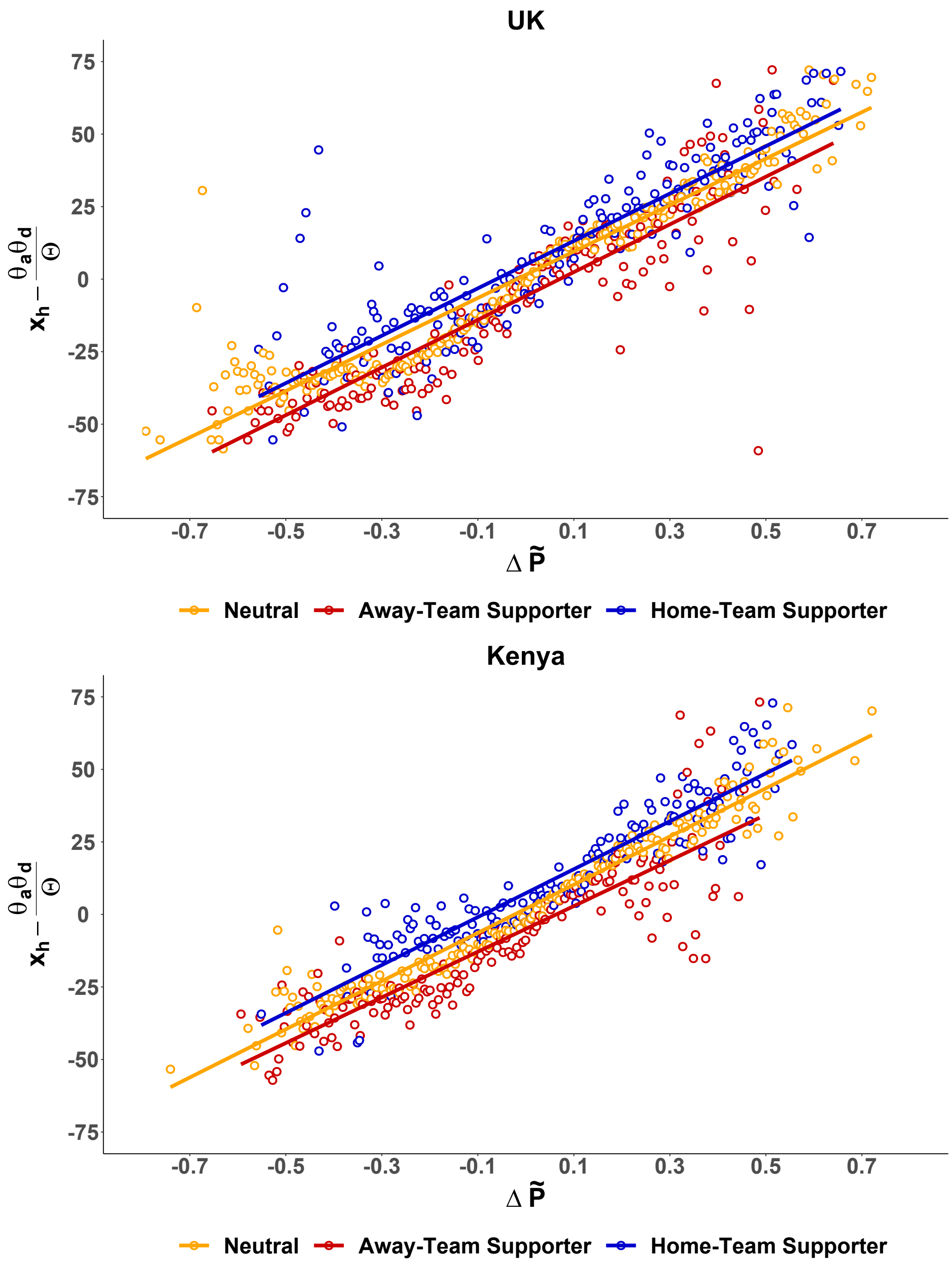
Notes: The top and bottom panels in the figure above summarize how survey participants in the UK and Kenya, respectively, allocate their betting budget across the three potential outcomes of a soccer match (home-win, draw, away-win). The budget allocation is categorized by whether participants support 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 A.1 and A.2.

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



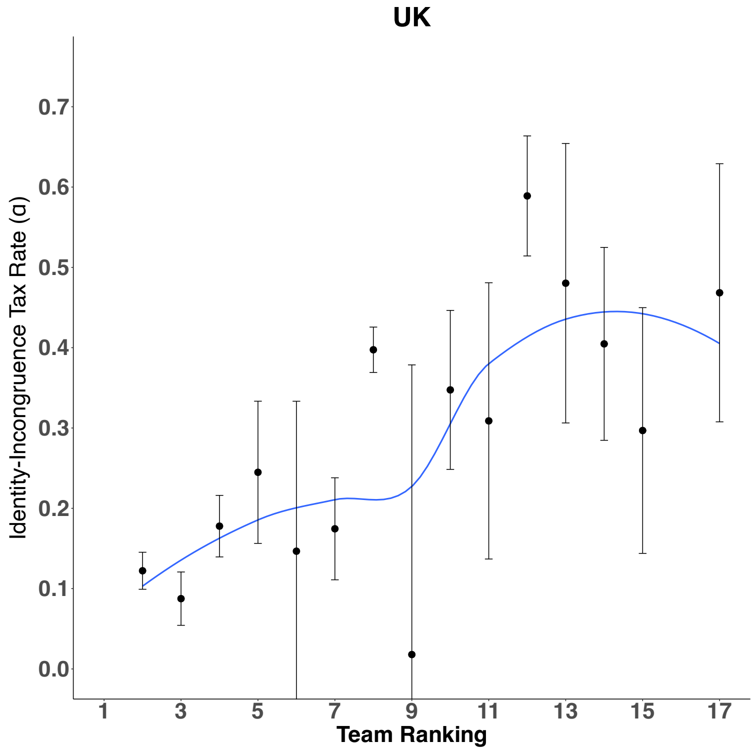
Notes: The top and bottom panels in the figure above show the CDFs of survey participants' self-reported beliefs about "home team winning" of Premier League soccer matches for the UK and Kenya, respectively. CDFs are first calculated for each match and then averaged across all matches.. The CDFs in each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

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



Notes: The panels in the figure above show a scatter plot of actual wagers on home-win (minus the ratio of odds provided) against the model-computed wedge term ($\Delta\tilde{P}$) 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 $\Delta\tilde{P}$, which captures the non-linear relationship of how beliefs and odds affect the amount wagered on a potential outcome.

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 participants in the UK. The bars indicate the range of +/-1 standard errors.

Tables

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 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.

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. 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 $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

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. 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 $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

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. 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 *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

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

	Dependent Variable: $x - \frac{\theta_d \theta_j}{\Theta} - r^{-1}\Delta\hat{P}$			
	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) and (2)) and for Kenya (columns (3) and (4)). In each column we regress wagers on home-win (or away-win) x_j minus $\frac{\theta_d \theta_j}{\Theta} - r_i^{-1}\Delta\hat{P}$ 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. 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 $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Semi-structural Portfolio Allocation (Heterogeneous Risk Preferences & Bookmaker Beliefs)

	Dependent Variable: $x - \frac{\theta_d \theta_j}{\Theta} - r^{-1}\Delta P^{BM}$			
	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. 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 $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7: 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 $^*p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 8: Counterfactual Simulations: Model Predicted Bets on Supported Team Win

	Counterfactual wager for equal odds across outcomes		
	Neutral Beliefs	Supporter Beliefs	Supporter Beliefs + Identity
	$\bar{x}^N(p^N, r)$	$\bar{x}^N(p^S, r)$	$\bar{x}^S(p^S, r)$
UK			
Bet (% of budget)	41.90%	45.05%	52.14%
%Δ relative to Neutral	-	7.52 %	24.44 %
Kenya			
Bet (% of budget)	49.50%	54.22%	60.17%
%Δ relative to Neutral	-	9.52%	21.54 %

Notes: This table simulates counterfactual wagers on supported team winning. The odds for home-team winning, drawing and away-team winning are set to be equal. 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 (3) and (5). They are averaged to team-level first for home-team supporters and away-team supporters separately, before we take the average of the two. Finally, we take the average for teams with over 20 supporters. Heterogeneous r and the corresponding α are applied. The second and third columns in the second and fourth rows are the percentage changes relative to the first columns. The simulations are calculated separately for the UK and Kenya.

Table 9: Heterogeneity in Identity-Incongruence Tax Rate

	Premier League Ranking	α	Standard Error
Change in α if supported team won their previous match	-	-0.096***	0.037
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	-	-0.002	0.646
Norwich City	-	0.686***	0.055
Watford	-	0.322	0.318

Notes: This table reports estimates of (1) the identity-incongruence tax rate for individual teams' ranking as ranked by their final position at the end of the recent prior Premier League season (2020/21) and (2) the change in identity tax if a team won their previous match. The estimates allow for individual-specific beliefs and risk preferences. The estimating equation corresponds to the first-order conditions (equations (5) and (6)), and individual and match-outcome fixed effects are accounted for in our estimation. The standard errors are computed via the delta method. The significance levels of our estimates are depicted as $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Appendix A Proof of Proposition 1

Proposition 1 If $\alpha_a, \alpha_d \geq 0$, and a fan is a supporter of the home team, then her optimal portfolio allocation is such that $x^S_{i,h} \geq x^N_{i,h}$.

Proof: Given:

- $u : R \to R$ is strictly increasing and strictly concave. Therefore, $u' > 0$,and $u'' < 0$.
- A neutral investor maximizes U^N_i

$$\text{Max}_{x_{i,h},x_{i,a}} U^N_i = p_{i,h}u(\theta_h x_{i,h}) + p_{i,a}u(\theta_a x_{i,a}) + p_{i,d}u(\theta_d x_{i,d}).$$

- A supporter investor maximizes U^S_i

$$\text{Max}_{x_{i,h},x_{i,a}} U^S_i = p_{i,h}u(\theta_h x_{i,h}) + (1 - \alpha_a)p_{i,a}u(\theta_a x_{i,a}) + (1 - \alpha_d)p_{i,d}u(\theta_d x_{i,d}).$$

- Both neutral and supporter investor have the same budget constraint

$$x_{i,h} + x_{i,a} + x_{i,d} = B$$

- Distortion due to identity concerns: $0 \leq \alpha_a, \alpha_d \leq 1$

We want to show that: $x^S_{i,h} \geq x^N_{i,h}$.

For a neutral investor, the FOCs for an interior solution are given by the equalization of the marginal utility per dollar spent on each outcome:

$$p_{i,h}u'(\theta_h x_{i,h})\theta_h = p_{i,a}u'(\theta_a x_{i,a})\theta_a = p_{i,d}u'(\theta_d x_{i,d})\theta_d. \qquad \text{FOC (1)}$$

For a supporter, due to the discount factors $(1 - \alpha_a)$ and $(1 - \alpha_d)$, the FOCs become:

$$p_{i,h}u'(\theta_h x_{i,h})\theta_h = (1 - \alpha_a)p_{i,a}u'(\theta_a x_{i,a})\theta_a = (1 - \alpha_d)p_{i,d}u'(\theta_d x_{i,d})\theta_d. \qquad \text{FOC (2)}$$

Comparing FOC (1) for neutrals and FOC (2) for supporters, we can see that the marginal utility per dollar spent on the away team and draw for the supporter is reduced by the factors $(1 - \alpha_a)$ and $(1 - \alpha_d)$ respectively.

Since $u'' < 0$, for the supporter to maintain the equality in FOC (2) with the presence of the α factors, they must increase $x_{i,h}$ such that:

$$p_{i,h}u'(\theta_h x^S_{i,h})\theta_h < p_{i,h}u'(\theta_h x^N_{i,h})\theta_h.$$

The only way this inequality can hold is if:

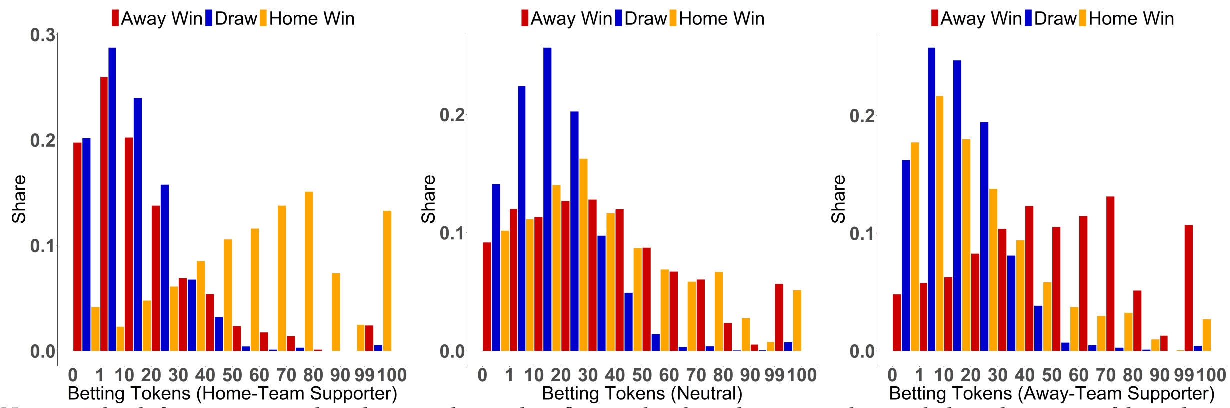
$$x^S_{i,h} > x^N_{i,h}$$

Thus, it must be that the supporter allocates strictly more to the home win compared to the neutral investor when the identity concerns are present. QED.

Appendix B Supplementary tables and figures

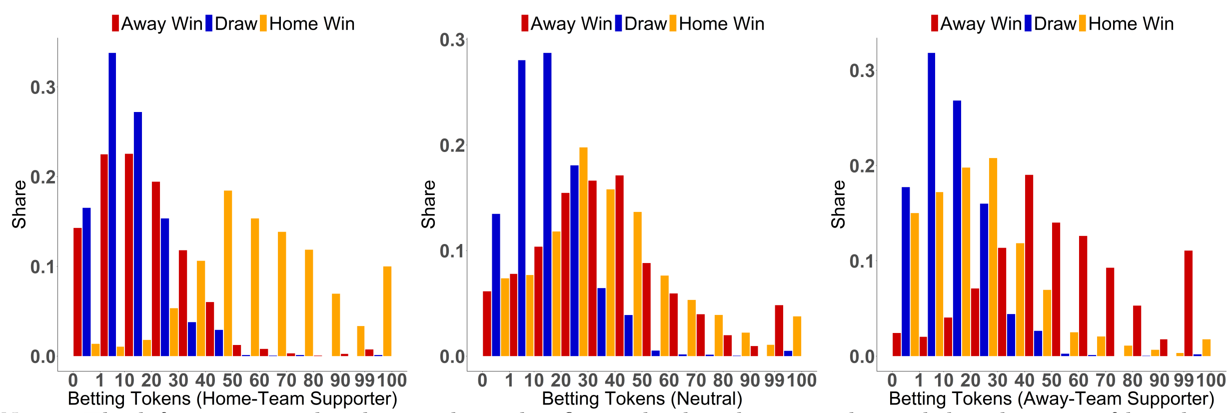
B.1 Figures

Figure A.1: Portfolio Allocations (UK): Unconditional Distributions



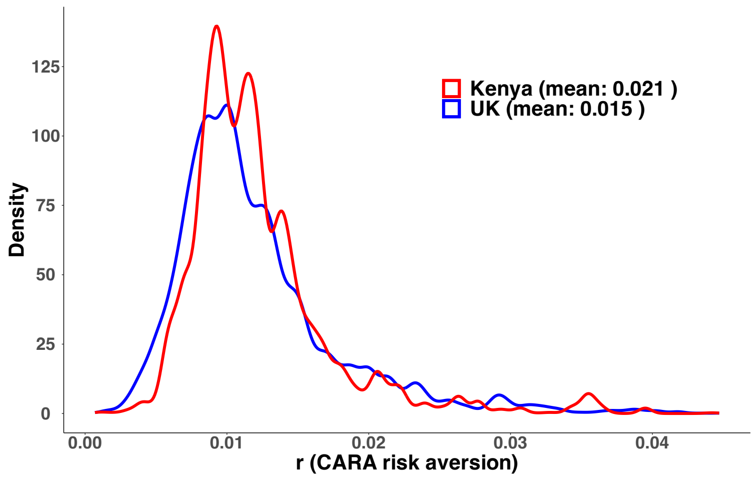
Notes: The left, center, and right panels in this figure display the unconditional distributions of how home supporters, neutrals, and away supporters in the UK allocate their betting budget across the three potential outcomes of a soccer match (home-win in orange, draw in blue, away-win in red).

Figure A.2: Portfolio Allocations (Kenya): Unconditional Distributions



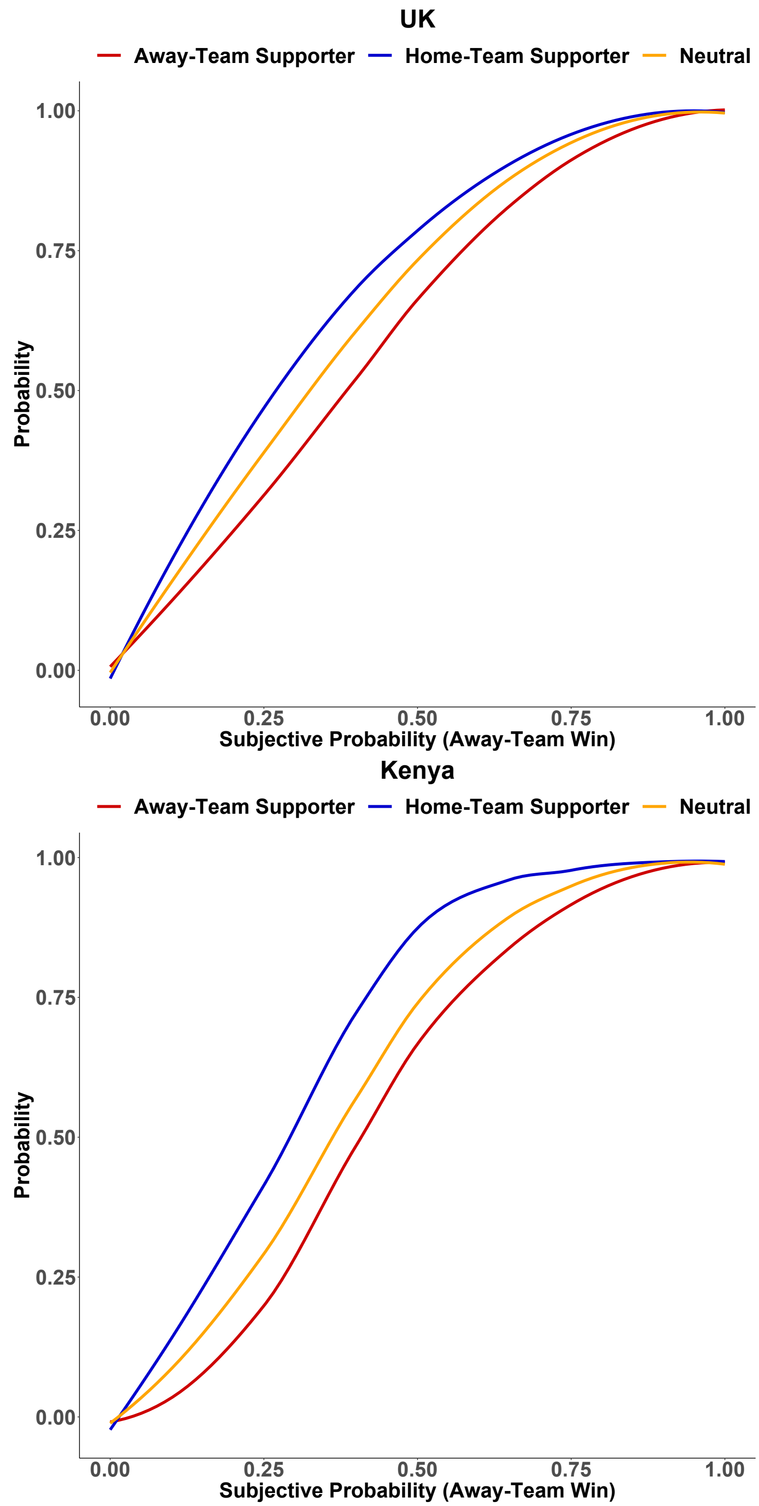
Notes: The left, center, and right panels in this figure display the unconditional distributions of how home supporters, neutrals, and away supporters in Kenya allocate their betting budget across the three potential outcomes of a soccer match (home-win in orange, draw in blue, away-win in red).

Figure A.3: Distribution of Risk Preferences



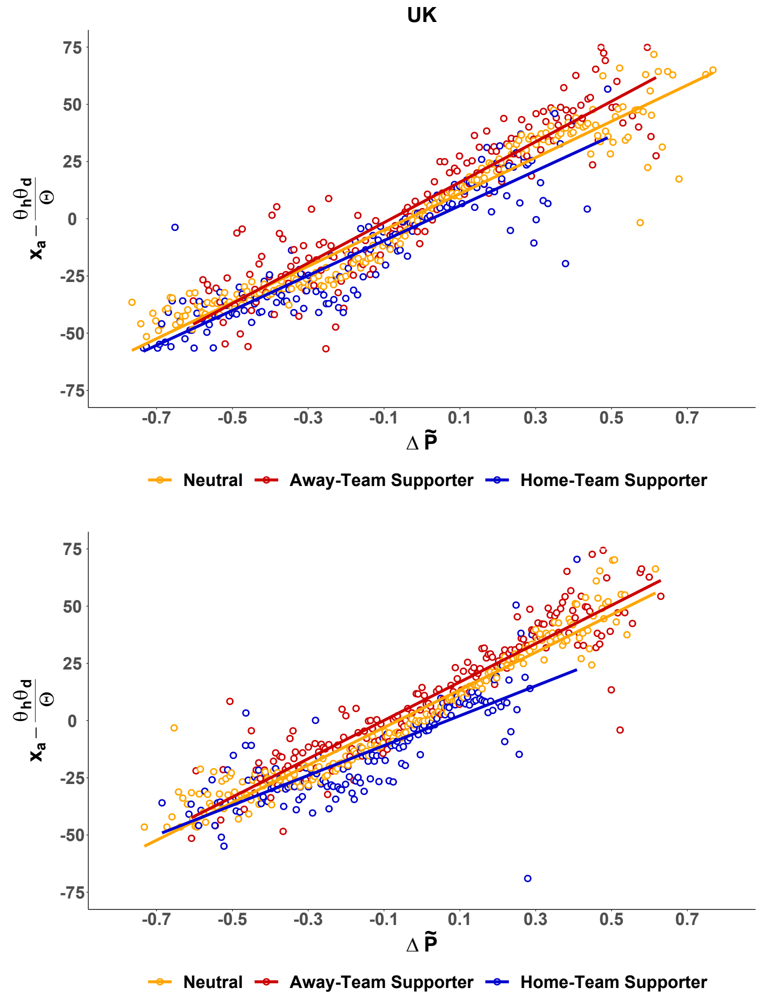
Notes: This figure presents the distribution of individual risk aversion estimates under the assumption of CARA and based on equation (8), separately for participants in the UK (blue) and Kenya (red).

Figure A.4: Subjective Beliefs (CDFs): Away-Team Win



Notes: The top and bottom panels in the figure 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 A.5: 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.

B.2 Tables

Table A.1: 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 A.2: Counterfactual Simulation at Team Level (UK Sample)

Team	Rank	α	Neutral Beliefs $\bar{x}^N(p^N, r)$	Supporter Beliefs $\bar{x}^N(p^S, r)$	Supporter Beliefs + Identity $\bar{x}^S(p^S, r)$	N	$\%\Delta_1$	$\%\Delta_2$
Man City	1	-	-	-	-	-	-	-
Man United	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	-	α >0	32.14	35.88	44.99	103.0	11.65	40.00
Average (N>20)	-	α >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 (3) and (5). 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 $\%\Delta_1 = (\bar{x}^N(p^S, r) - \bar{x}^N(p^N, r))/\bar{x}^N(p^N, r)$ and $\%\Delta_2 = (\bar{x}^S(p^S, r) - \bar{x}^N(p^N, r))/\bar{x}^N(p^N, r)$, respectively. The simulation is calculated only for the UK.

B.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 $z \in \{-2.5, -1.5, -.5, 0, +.5, +1.5, +2.5\}$,[20] where

$$z = p_a^0 * \theta_a^z - p_h^0 * \theta_h^z \tag{10}$$

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 (10) above, we get

$$p_a^0 * \frac{1}{p_a^0 - x^z} - p_h^0 * \frac{1}{p_h^0 + x^z} = z \tag{11}$$

or in terms of bookmakers' benchmark odds:

$$\frac{1}{\theta_a^0} * \frac{1}{\frac{1}{\theta_a^0} - x^z} - \frac{1}{\theta_h^0} * \frac{1}{\frac{1}{\theta_h^0} + x^z} = z \tag{12}$$

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

$$x = -\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z} + \sqrt{\left(\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z}\right)^2 + \frac{1}{\theta_a^0\theta_h^0}} \tag{13}$$

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

$$x = -\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z} - \sqrt{\left(\frac{\theta_h^0 + \theta_a^0 + z(\theta_a^0 - \theta_h^0)}{2\theta_a^0\theta_h^0 z}\right)^2 + \frac{1}{\theta_a^0\theta_h^0}} \tag{14}$$

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

Table A.3: Examples of Randomized Odds θ_h^z, θ_a^z for a Set of Benchmark Odds and Wedges z

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