Arbitrage in the Binary Option Market: Distinguishing Behavioral Biases

Aaron Goodman* Indira Puri*
MIT MIT

January 27, 2021

Abstract

In the first empirical analysis of the binary option market, we show that U.S. retail traders forgo clear arbitrage opportunities by purchasing binary options when strictly dominant portfolios of traditional call options are available at lower prices. Using a yearlong sample of binary option trades, we find that 19% of S&P index, 21% of gold, and 25% of silver trades violate our no-arbitrage condition. The amount of money lost is large, as buyers of binary options on average lose about a third of the contract price by forgoing the dominating call option portfolio. After rejecting standard institutional justifications for the existence of arbitrage, including random price volatility and various forms of trading costs, we examine possible behavioral explanations. We show that our results cannot be explained by canonical behavioral models such as prospect theory or cumulative prospect theory. Instead, we rationalize our findings with a novel behavioral model in which investors prefer simple binary lotteries to more complicated sets of outcomes. An online survey of binary option traders supplements our analysis of market data, providing direct evidence that a "preference for simplicity" is more common among these traders than prospect theory preferences.

*Contact: agoodm@mit.edu, ipuri@mit.edu. We thank participants at the 2021 AEA Meetings and the 2020 Economic Science Association Global Meetings for valuable comments. We are also grateful to Jaroslav Borovicka, Darrell Duffie, Drew Fudenberg, Basil Halperin, Adam Harris, Lisa Ho, Abhimanyu Mukerji, Jonathan Parker, Paul Pfleiderer, Jim Poterba, Charlie Rafkin, Jóse Scheinkman, Larry Schmidt, Amit Seru, John Sturm, Lawrence Summers, Adrian Verdelhan, Haoxiang Zhu for their feedback. Both authors acknowledge financial support from the George and Obie Shultz Fund at MIT and the National Science Foundation Graduate Research Fellowship Program under Grant No. 1122374. Puri acknowledges financial support from the Paul and Daisy Soros Fellowship for New Americans. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or other funding sources. This paper has also been circulated under the title 'Overpaying for Binary Options: Preference for Simplicity in Retail Markets.'

1 Introduction

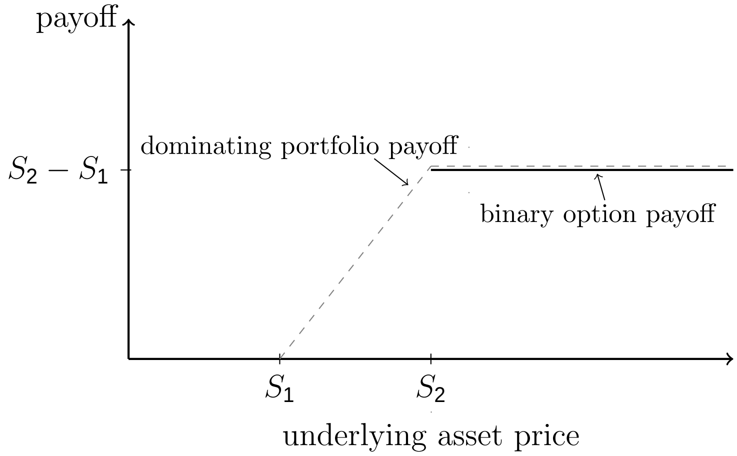
Despite their short history, binary options have generated no shortage of controversy. Since being introduced in 2008, these nonstandard derivative contracts – which pay a fixed amount if the buyer's bet is correct and zero otherwise – have caused such large losses among retail investors that they have been temporarily or permanently banned by the European Securities and Markets Authority (ESMA, 2019) and national governments in Israel (Weinglass, 2019), Denmark (Finanstilsynet, 2019), and the United Kingdom (FCA, 2019). Heightened public attention, however, has not produced a clear understanding of the factors responsible for retail traders' losses. Do they arise solely from isolated instances of fraud in European over-the-counter markets, or are they caused by fundamental behavioral biases that operate even at heavily regulated, price-transparent exchanges in the U.S.?

In a first empirical study of the binary option market, we find that retail binary option traders are losing significant amounts of money at regulated exchanges, and that these losses arise from the violation of a clear no-arbitrage condition relating binary option prices to the prices of traditional call options. After documenting these arbitrage opportunities, we ask what drives retail investors to forgo them. Our empirical strategy allows us to reject standard behavioral explanations such as prospect theory. We instead rationalize our findings with simplicity theory, a novel behavioral model introduced in Puri (2020), which posits that agents prefer lotteries with fewer possible outcomes. Because it accounts for types of suboptimal behavior that cannot be explained by prospect theory or other canonical behavioral models, the "preference for simplicity" we document among retail traders has important implications for studying and regulating households' behavior in other risky financial settings. Our empirical framework, which can distinguish among various forms of behavioral biases, may also be of independent interest.

Figure 1 summarizes the logic behind our empirical analysis. Because binary options give the holder a fixed payoff amount, a buy-sell portfolio of call options at nearby strike prices yields a payoff profile that is higher in every state of the world. Any well-known theory of choice under risk – including expected utility theory, cumulative prospect theory, and as we show, even prospect theory – predicts that willingness to pay for the strictly dominant call option portfolio should exceed willingness to pay for the binary option.

Gathering data for three different asset classes (S&P 500 index, gold, and silver options), we show that this no-arbitrage condition is often violated. For 19% of S&P, 21% of gold, and 25% of silver trades in our year-long sample of transaction data from the Nadex binary option exchange, we can find a portfolio of call options at the CME Globex exchange (using price quotes within 10 minutes of the binary option trade) that provides a strictly dominant payoff

Figure 1: A Dominated Binary Option



Note: We test whether the price of the dominating call option portfolio exceeds the price of the binary option.

profile and costs less. The resulting arbitrage price differences – which account for explicit trading fees at both exchanges – are large. Buyers of S&P, gold, and silver binary options on average lose 34%, 26%, and 38% of the binary option price by forgoing the dominating call option portfolio. Retail traders during our sample period thus left significant amounts of money on the table, often choosing to purchase binary options when cheaper, strictly superior alternatives were available.

After documenting frequent and large arbitrage opportunities, we show that they cannot be fully explained by the factors usually invoked to justify the existence of arbitrage in financial markets. In particular, we test and reject random price volatility, explicit trading fees, implicit trading costs arising from collateral requirements and liquidity differences, and differential investor knowledge as complete explanations for our results.

To assess the role of random price volatility, we compute arbitrage rates when the binary option is *dominating*, rather than *dominated*. If random price noise drives the results, then it should operate equally in both directions, meaning that arbitrage rates should be similar when the binary option is the better product. We show that this is not the case, as arbitrage rates are significantly smaller – both statistically and numerically – when the binary option is dominating, implying that random price noise cannot explain our results.

Our price comparisons between Nadex and CME control for explicit trading fees at both venues. Differences between the Nadex and CME exchanges in collateral requirements or market liquidity might create implicit trading costs, but we reject these possibilities as well. Because all positions that we analyze in this paper have weakly positive payoff profiles, neither Nadex nor CME require collateral to be posted in excess of the original trade price. Even so, we recognize that CME brokers may nonetheless require collateral and that the sequential execution of trades may create temporary loss exposure. For completeness, we thus test whether arbitrage opportunities are more frequent during periods of higher market volatility at CME (when CME collateral requirements are likely to be more stringent). The relationship between CME market volatility and arbitrage rates is too weak to play a substantial role in explaining our results. Liquidity considerations cannot account for our findings, either, since investors should be willing to pay a premium to trade in the more liquid CME market. Instead, we find that dominant CME portfolios often cost *less* than dominated Nadex binary options.

We also consider the possibility that Nadex's marketing efforts lead retail investors to be asymmetrically informed. If inexperienced investors without knowledge of CME's traditional option contracts are induced to trade binary options by Nadex's retail-focused advertisements, then they would not be in a position to compare price quotes between the Nadex and CME platforms and would not be able to act on the arbitrage opportunities we document. Since Nadex-related online search activity is likely to increase when targeted marketing campaigns succeed in driving new, inexperienced traders to Nadex's website, we test whether arbitrage rates are correlated with measures of Google search activity for the word "Nadex." We find no evidence of a positive correlation, suggesting that marketing-driven information distortions are not responsible for the observed arbitrage opportunities.

Having exhausted these standard institutional explanations, we next turn to behavioral theories. Rather than simply ascribing our results to behavioral biases, we use the existence of arbitrage to test and distinguish among theories of investor behavior. We prove mathematically that neither expected utility theory, nor prospect theory, nor cumulative prospect theory can explain the arbitrage opportunities arising from the mispricing of dominated binary options; indeed, any decision theory that respects first-order stochastic dominance is inconsistent with our results. We instead rationalize our findings with simplicity theory, a novel behavioral model introduced in Puri (2020). The theory is motivated by experimental studies and psychology literature showing that individuals are "complexity averse," with the "complexity" of a lottery depending in part on the number of possible outcomes. Using the formal theory and testable application to binary options from Puri (2020) as a basis for our empirical work, we show that a preference for simplicity can easily generate the observed arbitrage opportunities, with the relative simplicity of two-outcome binary option lotteries offsetting the price advantage of dominating many-outcome call option portfolios.

To supplement our trading analysis, we implement a survey of binary option traders that allows us to pose explicit lottery-choice questions and distinguish between different behavioral biases. In our sample, a plurality of binary option traders display a preference for simplicity in their choices, and simplicity preferences are substantially more common than prospect theory or cumulative prospect theory preferences. The survey results thus provide direct evidence that a preference for simplicity among retail traders drives the observed mispricing and arbitrage in the binary option market.

Finally, we append questions to our preference-elicitation survey in order to gauge the institutional knowledge of binary option traders. In ascribing binary option traders' behavior to simplicity theory, we assume that they are able to trade traditional options and choose to forgo them. The majority of respondents indicate that they are aware of traditional options and have brokerage accounts that grant trading access to CME, suggesting that this assumption is correct and confirming the result of our Google-search test for asymmetric investor knowledge. We are led to conclude that our results are not due to deficient institutional knowledge or limited access among retail binary option traders.

The rest of the paper is organized as follows. After a brief discussion of related literature, Section 2 provides an institutional overview of the binary option market. Section 3 derives the no-arbitrage condition that relates binary option prices to call option prices and outlines the empirical strategy we use to test this condition. Section 4 describes our data sources and the algorithm that we use to match Nadex binary option trades to dominating portfolios of CME call options. Section 5 presents the main arbitrage results, and Section 6 examines whether these results can be explained by standard justifications for arbitrage or canonical behavioral models. Section 7 discusses simplicity theory, which rationalizes our empirical results. Section 8 presents the survey results that we use to test directly for simplicity preferences and assess binary option traders' institutional knowledge. Section 9 concludes.

1.1 Related Literature

Our work contributes to three strands of literature: first, that on retail investors and their preferences; second, to a literature on distinguishing behavioral biases; and third, to a literature on complexity in financial markets. There is a large literature arguing that investors are susceptible to behavioral phenomena such as prospect-theory-like behavior (Barberis et al., 2016; Olsen, 1997), inattention (Bordalo et al., 2013; Graham and Kumar, 2006; Gabaix, 2019), and default effects (Carroll et al., 2009; Benartzi and Thaler, 2007). Relative to this literature, our paper builds by providing a new dimension on which retail investors may act behaviorally. We prove that prospect-theory-like behavior on the part of investors cannot explain our results, and show that canonical behavioral models are not sufficient to fully explain retail investors' choices over risky products.

There is a relatively small, experimental literature on tests that can rigorously rule out prospect theory (Andreoni and Sprenger, 2011; Fehr-Duda and Epper, 2012; Bernheim and Sprenger, 2020). We innovate relative to this literature by working in a real-world financial setting, and introducing a test which can rule out prospect theory preferences using prices. Our test also provides broad coverage: it rules out not only prospect theory but also any decision theory that respects dominance, as many do.

From our arbitrage test and survey data, we posit that our results can be explained by investors preferring simplicity (Puri, 2020), as defined formally in Section 7. Our paper therefore also contributes to the literature on complexity in financial markets (Carlin et al., 2013; Sato, 2014; Brunnermeier and Oehmke, 2009; Célérier and Vallée, 2012). This literature reaches conflicting conclusions about whether investors prefer or dislike complexity; these conflicts arise partially because the definition of complexity varies from study to study, each using an ad-hoc notion of complexity targeted for the specific setting of the paper. Our contribution to this literature is to use a precise mathematical definition of complexity which may be used in other domains.

While we are not aware of prior empirical studies of the binary option market, our work does relate to the literature on prediction markets, where the the traded contracts (e.g., those that pay $1 if a particular candidate wins the next presidential election) are similar in structure to fixed-payoff binary options. Our approach, however, is conceptually distinct. Whereas the prediction market literature assumes that traders act optimally and uses contract prices to infer the market's beliefs about event probabilities (Wolfers and Zitzewitz, 2004; Rhode and Strumpf, 2004; Snowberg et al., 2007), we derive a no-arbitrage condition and use data to determine whether traders adhere to it.

2 The Binary Option Market

Our binary option data come from Nadex, an online derivatives exchange that caters primarily to small retail investors.[1] Binary options are the most heavily traded of the three product types offered on the exchange,[2] and are marketed explicitly as a simple, easily

[1]See the exchange's website at `http://www.nadex.com`, and press articles surrounding its launch (under its original name of HedgeStreet.com), such as Business Wire (2008). The only other centralized trading venue for binary options in the U.S. is the Chicago Board Options Exchange; the CBOE binary option market is significantly smaller (both in terms of trading volume and the number of sponsored asset classes) than the Nadex market.

[2]The other product types are "touch brackets" and "call spreads," which differ slightly from binary options but also serve the purpose of simplifying and reducing the support of trading lotteries relative to traditional derivatives. In our year of trading data from Nadex, binary options account for about 94% of all trades with an identifiable product type.

understandable way to assume risk in financial and commodity markets. Instead of the continuous payoff profiles of traditional call and put options, binary options have only two possible payoff outcomes: if the underlying asset price is above the strike price at expiration, the seller pays the buyer a fixed amount; otherwise the buyer receives zero. Explaining the contract structure to prospective traders, Nadex describes binary options as "financial instrument[s] based on a simple yes or no question where the payoff is a fixed amount or nothing at all. This means binary options offer defined risk and clear outcomes on every trade."[3]

Despite its focus on small retail traders, the Nadex binary option market sees a substantial amount of trading activity. During our year of data between May 2018 and May 2019, Nadex executed over 4.2 million binary option trades. The aggregate market value of these trades (using trade prices at the time of execution) was over $513 million, and the aggregate notional value (with all Nadex binary options giving a $100 fixed payoff) was just over $1.04 billion. As discussed in Section 4, in carrying out our no-arbitrage test, we must focus on Nadex contracts for which there exist traditional options at CME with the same underlying asset and expiration date. As a result, our analysis sample is restricted to weekly and Friday daily contracts referencing S&P index, gold, and silver prices. Among this smaller analysis sample, our year of data contains 54,142 trades with a market value of $11.2 million and a notional value of $22.2 million.

According to a Commodity Futures Trading Commission report issued in July 2017, only two institutional investors had been licensed to practice as market makers at Nadex, and one of these firms is itself a subsidiary of Nadex's parent company. The same CFTC report noted that one of these two market makers was on one side of 99 percent of all trades executed during the period under study (with the Nadex-affiliated market maker taking part in 70 percent of all trades).[4] Retail investors trading on Nadex during our sample period[5] were thus transacting with institutional counterparties who faced very little competition in setting prices because so few institutional investors had been admitted to trade on the exchange. In effect, Nadex itself (through its affiliated market maker) was selling derivative contracts to retail investors, and other institutional investors who might have traded on arbitrage opportunities did not have access to the exchange. This lack of external arbitrageur capital would allow arbitrage opportunities to persist, and the monopolistic market makers profiting

[3]See Nadex (2020).

[4]See pg. 16 of the report Commodity Futures Trading Commission Division of Market Oversight (2017).

[5]Note that the CFTC report (Commodity Futures Trading Commission Division of Market Oversight, 2017) was issued in July 2017, while our sample period runs from May 2018 to May 2019. We thus do not know the exact number of market makers operating on Nadex during our sample period, but we do not expect the number to have increased substantially between July 2017 and May 2018.

from their trades with retail investors would have no incentive to move prices back within no-arbitrage bounds. Additionally, given the relatively small trading volume and modest amount of arbitrage profits available at the retail-focused Nadex exchange, we doubt that many sophisticated institutional investors would be willing to incur the fixed costs of setting up Nadex trading access, even if the exchange allowed them to do so. In Appendix A we outline a formal model in which a monopolistic Nadex market maker and a lack of external arbitrage capital allow arbitrage price differences to persist.

3 Empirical Strategy

Figure 1 illustrates the logic behind our main empirical analysis, which follows the logic of the test earlier proposed in Puri (2020). A binary option with a strike price of S_2 pays 0 to the buyer if the underlying asset price is below S_2 at contract expiration, and pays some fixed positive amount if the underlying asset price finishes above S_2. Graphically, this payoff profile is the solid horizontal line extending rightward from S_2.

Unlike the fixed payoff of a binary option, a traditional call option that settles in the money pays the buyer the difference between the underlying asset price and the contract's strike price. Buying a call option at strike price S_1 thus creates a payoff profile extending rightward from S_1 with a slope of 1, as shown by the dashed line in Figure 1. If the difference between the strike prices S_1 and S_2 is chosen to match the binary option's fixed payoff amount, this upward-sloping dashed line intersects with the solid binary option payoff exactly at S_2. Finally, if the buyer of the call option with strike price S_1 also sells a call option with strike price S_2, the losses from the latter exactly cancel the winnings from the former as the underlying asset price moves above S_2. The net payoff profile of this two-contract portfolio (buying a call at S_1 and selling a call at S_2) thus levels off and becomes identical to the horizontal binary option payoff for all underlying asset prices above S_2.

Comparing the payoff profiles in Figure 1, two facts are clear. First, the binary option induces a lottery with a smaller support: while the binary option can only pay 0 or $S_2 - S_1$, the portfolio of call options can pay anything between 0 and $S_2 - S_1$. Second, we can see that the payoff profile of the call option portfolio strictly dominates the payoff profile of the binary option. The payoffs are identical for all underlying asset prices below S_1 or above S_2, but the portfolio of call options pays off a strictly higher amount than the binary option for all underlying asset prices between S_1 and S_2. For the market to be free of arbitrage, there must be strictly positive state prices associated with states of the world where the underlying asset price is between S_1 and S_2, and the call option portfolio must cost more than the binary option. Formally, denoting the prices of a call option and a binary option with strike price i as P_i and B_i, respectively, we have:

$$\text{no arbitrage} \implies P_{\mathsf{S}_1} - P_{\mathsf{S}_2} > B_{\mathsf{S}_2}. \tag{1}$$

The goal of our empirical analysis is to determine how often condition (1) is violated. Using binary option data from Nadex and traditional call option data from CME, we match observed binary option trades to price quotes for call options around the same time. We can then learn the fraction of Nadex trades that took place when a dominating portfolio was available at a lower cost at CME.

4 Data and Trade-Matching Algorithm

4.1 Nadex Data

Nadex offers binary option trading on a range of underlying assets in equity, foreign exchange, and commodity markets. Because our strategy depends on the existence of comparable call options traded on CME in reference to the same underlying asset, during the same time period, and with the same expiration date, only a subset of the Nadex binary option markets are valid settings for our proposed empirical test. These considerations lead us to focus on the markets for S&P 500 index, gold, and silver binary options. The underlying asset price for Nadex S&P index options is the near-month CME E-Mini S&P 500 Index futures price, and for gold and silver options, it is the near-month COMEX futures price.

As discussed below, the CME call options we use are weekly options that expire on Friday afternoon. Nadex's weekly binary options also expire on Friday afternoon and are thus directly comparable to the CME call options. Nadex also hosts daily contracts that begin trading in the evening and expire the following afternoon. Nadex's Friday daily contracts expire at the same time on Friday afternoon as Nadex's weekly contracts, so they are also directly comparable to the CME weekly call options (albeit with a shorter trading life). To maximize our sample size, we consider both Nadex's weekly contracts and Friday daily contracts. One of the robustness checks below in Section 5 shows that our arbitrage results do not differ significantly between the weekly and Friday daily contracts.

We observe all trades in weekly and Friday daily contracts for the S&P, gold, and silver markets for one full year between May 17, 2018 and May 17, 2019. For each observed trade, the Nadex data show the execution time, the strike price, the transaction price, and the number of contracts traded. The payoff profile for buyers of these options is exactly as shown in Figure 1, with the fixed payoff amount equal to $100 for all contracts. There is also a $1 trading fee per contract, which we factor into our calculations when comparing Nadex transaction prices to CME price quotes.

4.2 CME Globex Data

Our traditional call option data come from CME Globex, a large online derivatives exchange that offers futures and option trading in a wide variety of underlying assets and serves institutional and professional investors as well as retail traders. The CME weekly options we use expire on Friday afternoon[6] and settle to the same underlying asset price (near-month CME E-Mini S&P 500 Index futures price for S&P index options, near-month COMEX futures price for gold and silver options) as the Nadex binary options described above.[7] We obtain top-of-book data from CME for the same yearlong sample period between May 17, 2018 and May 17, 2019. The CME data show changes in the top-of-book best bid and ask price quotes as well as executed trades. For each price quote or executed trade, we observe a time stamp, the strike price, the quoted or transaction price, and the number of contracts listed or traded. The side of the market (i.e., ask vs. bid) is indicated for changes in top-of-book price quotes, but not for executed trades.

We restrict the CME data to call options, which can be combined as shown in Figure 1 to create payoff profiles that dominate those of Nadex binary options with nearby strike prices. If these call options settle out of the money, buyers receive nothing; if they settle in the money, buyers receive the difference between the underlying asset price and the strike price, multiplied by the notional amount of the contract. The notional amounts of the S&P, gold, and silver options are $50, 100 troy ounces, and 5,000 troy ounces, respectively. CME charged a trading fee of $0.55 per contract for weekly S&P options and $1.45 per contract for weekly gold and silver options during our sample period, which we factor into our calculations when computing price differences.[8]

[6] Note that while both the Nadex binary options and the CME call options we study expire on Friday afternoon, they do not expire at exactly the same time. Nadex's weekly and Friday daily contracts in gold expire at 1:30pm ET, and those in silver expire at 1:25pm ET. CME's weekly gold and silver options both expire at 5:00pm ET. Since options that expire later are more valuable, this difference in expiration times only makes the CME portfolios we construct more appealing relative to Nadex binary options, and thus cannot explain arbitrage opportunities where dominating CME portfolios in gold and silver cost less than their corresponding binary options. As for the S&P index options, Nadex's weekly and Friday daily contracts expire at 4:15pm ET while CME's weekly options expire at 4:00pm ET. The difference in expiration times for the S&P index options thus works the other way and would tend to make Nadex binary options more valuable (all else equal) than corresponding CME portfolios. However, we expect any valuation difference arising from this 15-minute discrepancy to be small, and incapable of explaining the frequent and large arbitrage opportunities we document in Section 5.

[7] The CME exchange tickers for the S&P, gold, and silver options we use begin with "EW," "OG,", and "SO," respectively.

[8] We use the trading fees charged to non-members. This is a conservative assumption because fees charged to non-members are higher than fees charged to CME members, and we seek to find arbitrage opportunities where dominating CME portfolios cost less than their corresponding binary options.

4.3 Trade-Matching Algorithm

We carry out our empirical strategy by matching observed Nadex trades to CME price quotes and trades around the same time, then computing price differences to test for violations of the no-arbitrage condition (1). In particular, for each Nadex trade, we do the following:

- Subset CME call option quotes and trades to those that occurred in the 10 minutes before the Nadex trade. We do this in order to identify the CME prices (and any resulting arbitrage opportunities) that the buyer of the Nadex binary option passed up in the minutes before executing his or her trade.

- Further subset the CME quotes and trades to those at the two highest strike prices that are weakly less than the Nadex trade's strike price. Formally, denoting S_{nadex} as the Nadex trade's strike price and $\{S_i\}_{i=1}^{N}$ as the N unique strike prices among the CME quotes and trades, define:

$$S_{close} = \max\{S_i | S_i \leq S_{nadex}\}, \tag{2}$$

$$S_{far} = \max\{S_i | S_i < S_{close}\}. \tag{3}$$

- For both strike prices S_{close} and S_{far}, choose one CME quote or trade price to use in evaluating the no-arbitrage condition. Do so by using the following priority ordering:

  - First, prioritize price quotes that are on the correct side of the market, assuming that the trades used to construct our dominating CME portfolio would be taking liquidity. In other words, since our CME portfolio is constructed by buying at S_{far} and selling at S_{close}, prioritize quotes at S_{far} that are on the ask side of the market, and quotes at S_{close} that are on the bid side of the market. Note that the CME data indicates market side for price quotes but not for executed trades; we thus use executed trade prices only if we cannot find an appropriate price quote in our 10-minute window. We flag instances where we are forced to use a quote from the wrong side of the market or an executed trade price and exclude them in one of our robustness checks.

  - Second, prioritize quotes and trades that occurred closer in time to the Nadex trade.

- Finally, compute the difference between the Nadex trade price and the price of the dominating CME portfolio, using the CME quote or trade prices chosen in the above steps.

– The price of the CME portfolio is

$$P_{S_{close}} - P_{S_{far}} + 2F, \tag{4}$$

where P_i is the price of the CME option with strike price i and F is the CME trading fee per contract ($0.55 for S&P options and $1.45 for gold and silver options).

– The difference in strike prices $S_{far} - S_{close}$ determines the height at which the CME payoff profile levels off into a horizontal line (see Figure 2 below). Since the fixed payoff amount of Nadex binary options is always $100, we must scale the Nadex trade quantity to ensure that the Nadex and CME payoff profiles are the same at all underlying asset prices above S_{nadex}. The price of the scaled Nadex trade is thus

$$\frac{M * (S_{far} - S_{close})}{100} * (1 + B_{S_{nadex}}), \tag{5}$$

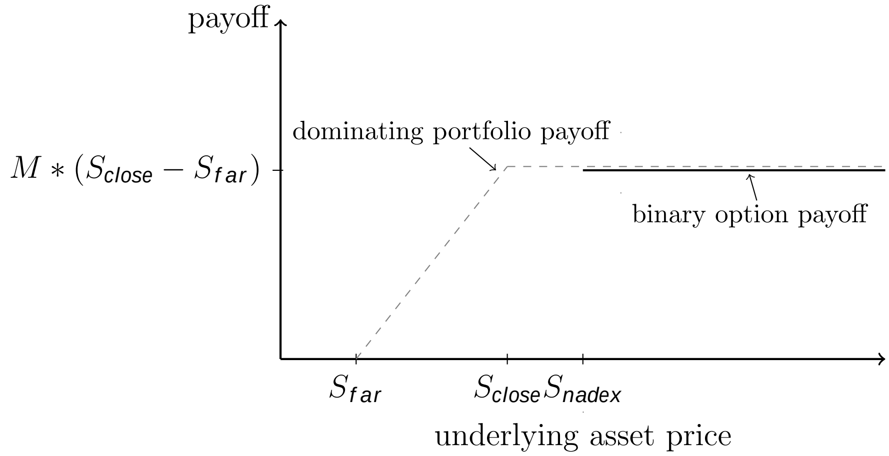
where M is the notional amount of the CME contract ($50 for S&P options, 100 troy ounces for gold, and 5,000 troy ounces for silver), B_i is the price of the Nadex binary option with strike price i, and the additional $1 term comes from Nadex's $1 trading fee per contract.

– Evaluating the no-arbitrage condition thus reduces to

$$P_{S_{far}} - P_{S_{close}} + 2F \lesseqgtr \frac{M * (S_{far} - S_{close})}{100} * (1 + B_{S_{nadex}}). \tag{6}$$

If the left-hand side of expression (6) is smaller, we conclude that the dominating CME portfolio costs less than the scaled Nadex trade and the no-arbitrage condition is violated. Note that the scale factor $\frac{M \ (S_{far} \ S_{close})}{100}$ is not always an integer, and trading a fractional number of Nadex contracts is not possible under exchange regulations. We thus adjust fractional scale factors down to the closest possible integer scale factor that preserves the CME portfolio's dominating status (though we conduct an additional robustness check in which we allow fractional scale factors). Note also that the observed trade quantity in the Nadex data is usually smaller than the scale factor, so we are implicitly assuming that the Nadex trader would have been willing to scale up his or her trade and purchase $\frac{M \ (S_{far} \ S_{close})}{100}$ contracts. We do not have to make this assumption when the observed trade quantity is actually greater than or equal to $\frac{M \ (S_{far} \ S_{close})}{100}$, so a final robustness check below limits to these cases.

Figure 2: Trade-Matching Algorithm



The payoff profiles compared by our trade-matching algorithm are shown in Figure 2. It is worth noting that nearest CME strike price S_{close} is generally not equal to the Nadex strike price S_{nadex} (Figure 2 illustrates such a case). The logic in evaluating the no-arbitrage condition remains exactly the same whether or not S_{close} is equal to S_{nadex}. The important difference is that when S_{close} does not equal S_{nadex}, or when S_{far} is further from S_{nadex}, the CME portfolio gives a strictly higher payoff than the Nadex contract over a larger range of underlying asset prices. The resulting arbitrage opportunity (if it exists) is thus more severe. When discussing arbitrage rates below, we disaggregate the results by the distance between the Nadex strike price and the matched CME strike prices.

4.4 Summary Statistics

Table 1 presents summary statistics for our sample of Nadex binary option trades. Together, the S&P, gold, and silver binary options provide a sample of 54,142 trades that we attempt to match to dominating CME portfolios using the algorithm described in Section 4.3. Of these 54,142 total trades, 25,502 are weekly contracts while the remaining 28,640 are Friday daily contracts. Nadex traders during our sample period put moderate amounts of money at risk: the average trade price for all three option types is around $50 per contract, and the average sizes of S&P, gold, and silver trades are 4.26, 3.77, and 2.10 contracts, respectively. Because the CME top-of-book data contain millions of price quotes and trades per day, our matching algorithm is successful most of the time, as we match 79% of S&P trades, 87% of gold trades, and 76% of silver trades to a dominating CME portfolio.

Aside from showing the overall success rate in matching Nadex trades to dominating CME

Table 1: Summary Statistics for Nadex Binary Option Trades

	S&P	Gold	Silver
Trade Price	51.33	46.11	46.37
	(24.06)	(24.74)	(26.94)
Quantity	4.26	3.77	2.10
	(13.98)	(10.45)	(3.28)
Matched to CME Portfolio	0.79	0.87	0.76
	(0.40)	(0.33)	(0.43)
Time Diff. to Close Strike (secs)	33.42	44.42	72.04
	(93.63)	(92.82)	(114.47)
Time Diff. to Far Strike (secs)	34.33	35.59	63.24
	(93.57)	(81.87)	(109.20)
Close Strike Distance	2.27	2.64	0.02
	(2.02)	(8.45)	(0.10)
Far Strike Distance	7.43	7.81	0.08
	(3.22)	(9.81)	(0.20)
Scale Factor	2.08	5.17	2.94
	(0.97)	(4.99)	(7.93)
Weekly Trades	19,572	4,875	1,055
Friday Daily Trades	24,673	3,101	866
Total Trades	44,245	7,976	1,921

Note: All variable names are self-explanatory except those concerning close and far strike prices and the scale factor. See Section 4.3 and Figure 2 for definitions of the close and far strike prices in the matched CME portfolio. The distance variables here give the difference between the Nadex strike price and the close/far CME strike prices (in index points for the S&P options, in dollars per troy ounce for the gold and silver options). The scale factor is the number of Nadex contracts that must be purchased in order to match the height of the CME portfolio's payoff profile; again see Section 4.3 and Figure 2 for discussion of the scale factor.

portfolios, Table 1 also summarizes how "close" the Nadex trades are to their matched CME portfolios, in terms of both time and strike price placement. The average time difference between a Nadex trade and the price quotes or trades in its matched CME portfolio is about 34 seconds for S&P options, about 40 seconds for gold options, and about 67 seconds for silver options (depending on whether the close or far strike price is considered). Our 10-minute cutoff for locating CME prices is thus rarely binding, and tightening this cutoff does not noticeably change the results. The matched CME portfolios also contain strike prices that on average are close to the Nadex strike price. However, the standard deviations of these strike price differences are fairly large, and Section 5 disaggregates the main arbitrage results based on these distances (with larger strike price distances indicating larger or more severe arbitrage opportunities). Finally, Table 1 summarizes the scale factor that we use to match the payoff profiles of Nadex options and their corresponding CME portfolios. As discussed in Section 4.3, to ensure that the payoff profiles of the Nadex option and CME portfolio level off at the same height, we must purchase $\frac{M\ (S_{close}\ \ S_{far})}{100}$ (rounded down to the nearest integer) Nadex contracts. The average scale factor is 2.08, 5.17, and 2.94 for S&P, gold, and silver options, respectively.

5 Arbitrage Results

Our matching algorithm produces, for each matched Nadex trade, the price of a dominating CME portfolio that was available just before the time of the Nadex trade. Cases where the dominating CME portfolio costs less than the (appropriately scaled) Nadex binary option constitute arbitrage opportunities. Figure 3 summarizes the distribution of these price differences, in both absolute dollar terms (CME portfolio price minus the scaled Nadex price) and relative terms (the absolute difference divided by the price of the scaled Nadex trade). In each case there is substantial mass to the left of zero, indicating that violations of the no-arbitrage condition arise often.

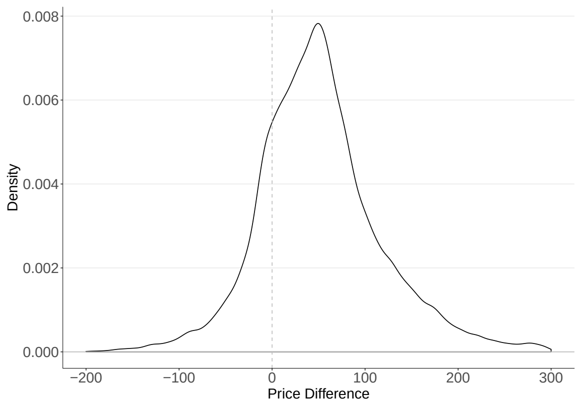
The main arbitrage results are presented in Table 2. The baseline results, which consider all Nadex trades that were successfully matched to a dominating CME portfolio, show that arbitrage occurs often. We find that 19% of the 33,676 matched S&P options, 21% of the 6,876 matched gold options, and 25% of the 1,355 matched silver options cost more than their dominating CME portfolio.[9]

Furthermore, conditional on arbitrage existing, the price differences between Nadex binary options and dominating CME portfolios are large. Because the distributions of price differences are generally right-skewed, we report means as well as medians. In the baseline sample the mean price differences for S&P, gold, and silver trades are $41.28, $51.01, and $28.30, while the median price differences are $21.90, $33.35, and $19.60, respectively. In relative terms, the price difference represents an average of 34%, 26%, and 38% of the binary option trade price for the S&P, gold, and silver asset classes, with the median figures slightly smaller. Summing these arbitrage price differences across trades, we find total arbitrage amounts during our one-year sample period of $265,500 for S&P contracts, $72,441 for gold contracts, and and $9,734 for silver contracts. These are of course relatively modest amounts, but in drawing conclusions about retail investor behavior, we are more interested in the

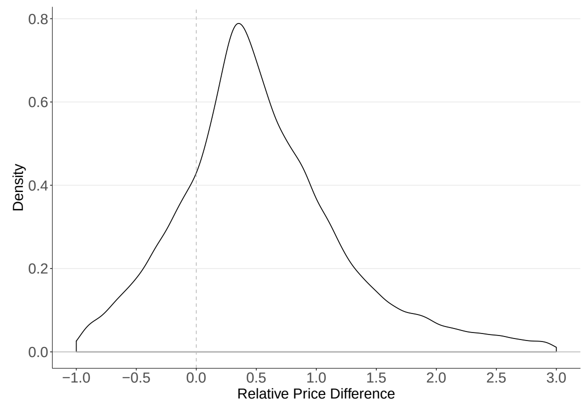
[9]For a small number of Nadex trades, the matched CME call option at the far strike price actually costs less than the matched CME call option at the close strike price (see again Figure 2). In these cases (which likely arise from outlying trade prices or errors in the CME data feed), arbitrage would exist regardless of the Nadex trade price. We therefore drop these trades from the sample before computing our arbitrage statistics.

Figure 3: Price Differences Between Dominating CME Portfolios and Nadex Binary Options

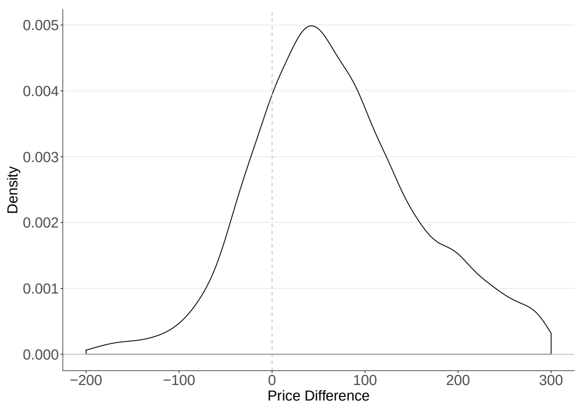
(a) S&P: Absolute Price Differences



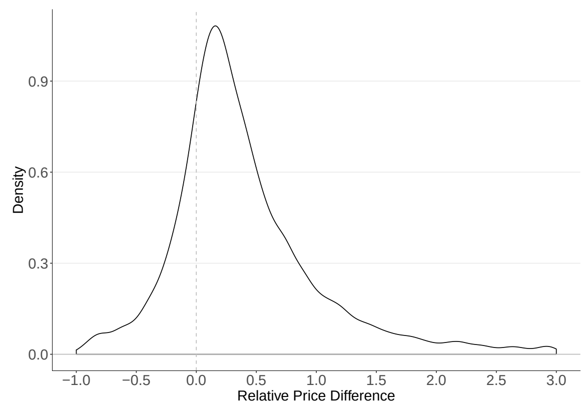
(b) S&P: Relative Price Differences



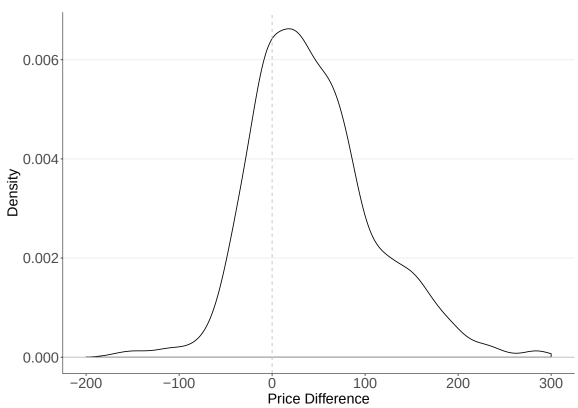
(c) Gold: Absolute Price Differences



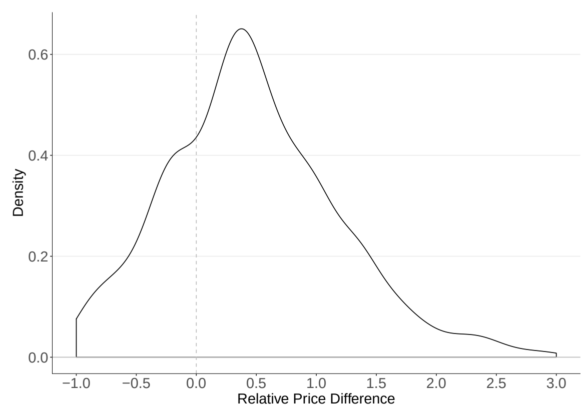
(d) Gold: Relative Price Differences



(e) Silver: Absolute Price Differences



(f) Silver: Relative Price Differences



Note: These graphs summarize the price differences between Nadex binary options and their matched CME portfolios. The plotted variable is the price of the matched CME portfolio minus the price of the Nadex binary option; negative values thus indicate arbitrage opportunities. The absolute price difference is the raw dollar value, and the relative price difference divides the absolute price difference by the price of the scaled Nadex trade (i.e., the Nadex trade price times the scale factor).

Table 2: Arbitrage Results: Baseline

	S&P	Gold	Silver
Arbitrage Rate	0.19	0.21	0.25
Arbitrage Size			
Mean	41.28	51.01	28.30
Median	21.90	33.35	19.60
Arbitrage Size/Price			
Mean	0.34	0.26	0.38
Median	0.29	0.18	0.30
Total Arbitrage Amount	265,500	72,441	9,734
Nadex Trades Considered	33,676	6,876	1,355

Table 3: Arbitrage Results: Robustness

	CME Prices on Correct Side			Weekly Trades Only			Scale Factor $\leq$ Quantity			Fractional Scale Factors		
	S&P (1)	Gold (2)	Silver (3)	S&P (4)	Gold (5)	Silver (6)	S&P (7)	Gold (8)	Silver (9)	S&P (10)	Gold (11)	Silver (12)
Arbitrage Rate	0.18	0.19	0.24	0.18	0.23	0.30	0.19	0.20	0.28	0.33	0.21	0.37
Arbitrage Size												
Mean	32.25	47.55	28.53	34.85	43.97	25.44	36.22	44.55	27.95	44.26	51.01	36.82
Median	20.90	32.10	20.60	20.40	30.85	17.35	22.90	28.35	18.60	26.40	33.35	25.23
Arbitrage Size/Price												
Mean	0.32	0.23	0.36	0.32	0.24	0.38	0.33	0.24	0.38	0.31	0.26	0.38
Median	0.27	0.17	0.29	0.28	0.17	0.30	0.28	0.15	0.30	0.26	0.18	0.35
Total Arbitrage Amount	194,495	61,620	8,760	100,288	45,817	6,157	88,952	8,331	3,326	488,530	72,441	18,371
Nadex Trades Considered	32,658	6,673	1,281	15,927	4,437	820	12,962	945	423	33,676	6,876	1,355

Note (to Tables 2-3): Only Nadex trades that were successfully matched to a dominating CME portfolio are considered. The arbitrage rate is the fraction of the relevant Nadex options that cost more than their dominating CME portfolio. Arbitrage size is the difference between the scaled Nadex trade price (i.e., the Nadex trade price times the scale factor) and the cost of the dominating CME portfolio, conditional on the Nadex option costing more and arbitrage existing. Arbitrage size/price normalizes arbitrage size relative to the price of the scaled Nadex trade. Total arbitrage amount is the sum of arbitrage sizes across all trades. See text for discussion of the sample splitting employed in each column grouping.

frequency with which no-arbitrage conditions are violated (given by the arbitrage rate) than in the size of this particular market. We also note that we focus on weekly and Friday daily contracts in these three asset classes because of current data availability and the requirement that comparable call options be traded on CME, but in doing so we miss most of the larger Nadex binary option market (see Section 2).

Table 3 shows four robustness checks that were introduced during the discussion of the matching algorithm in Section 4.3. Columns 1-3 restrict to cases where we are able to find CME price quotes on the correct side of the market – i.e., cases where the price quote at the far strike price comes from the ask side of the market and the price quote at the close strike price comes from the bid side of the market. Since we are almost always able to find CME price quotes on the correct side of the market, the sample sizes only decrease by a small amount and the results are largely unchanged. Columns 4-6 drop Nadex's Friday daily options and consider only weekly options; the results do not change significantly. Columns 7-9 limit to trades where the observed trade quantity in the Nadex data is at least as large as the scale factor: in these cases, our assumption that the Nadex trader would have been willing to purchase the scaled-up number of Nadex contracts is obviously true. The sample sizes decrease significantly but the results are again essentially unchanged relative to the baseline. Finally, columns 10-12 allow fractional scale factors (recall that these results are of secondary importance because Nadex exchange rules do not allow trades of non-integer numbers of contracts). Relaxing the constraint on integer scale factors increases arbitrage rates but does not have much of an effect on conditional arbitrage sizes.

Finally, we disaggregate the results by strike price distance. Again referring to Figure 2, the relative placements of S_{nadex}, S_{close}, and S_{far} differ between matched Nadex option-CME portfolio combinations. The further apart the strike prices, the larger the range of underlying asset prices for which the CME portfolio provides a strictly higher payoff than the Nadex option, and the larger or more severe the potential arbitrage opportunity.

Tables 4-6 split the sample of matched Nadex trades by the distance between the Nadex strike price and the far strike price in the dominating CME portfolio. The strike price distances are in units of S&P index points for the S&P options and in units of dollars per troy ounce for the gold and silver options. Each column in Tables 4 and 5 represents an interval of strike price distances; because strike price distances for silver options take on a very small number of unique values in our sample, each column in Table 6 represents a single strike price distance. For both S&P and gold options, arbitrage rates decrease monotonically as the strike price distance grows (from a maximum of 21% to a minimum of 16% for S&P options; from 27% to 15% for gold options). We are only able to split the silver options into two meaningful subsamples, but the arbitrage rate does decrease from 32% to 20% when the

Table 4: S&P Arbitrage Results, by Distance to Far CME Strike Price

	Strike Price Distance (index points)				
	[5, 6)	[6, 7)	[7, 8)	[8, 9)	[9, 10]
Arbitrage Rate	0.21	0.21	0.19	0.16	0.16
Arbitrage Size					
Mean	33.02	30.08	34.15	30.29	42.32
Median	21.40	20.90	21.90	18.40	25.40
Arbitrage Size/Price					
Mean	0.35	0.31	0.35	0.31	0.35
Median	0.30	0.26	0.30	0.23	0.31
Nadex Trades Considered	7,204	6,804	6,405	6,408	6,508

Note (to Tables 4-6): Only Nadex trades that were successfully matched to a dominating CME portfolio are considered. See notes to Table 3 for definitions of the arbitrage measures. Nadex trades are disaggregated by the distance between the Nadex strike price and the far strike price in the dominating CME portfolio (see Figure 2 for a graphical depiction of strike price distances). A small number of outlying Nadex trades with strike price distances larger than the maximum values shown in the tables are excluded.

Table 5: Gold Arbitrage Results, by Distance to Far CME Strike Price

	Strike Price Distance (dollars/troy ounce)				
	[5, 6)	[6, 7)	[7, 8)	[8, 9)	[9, 10]
Arbitrage Rate	0.27	0.21	0.20	0.17	0.15
Arbitrage Size					
Mean	52.60	55.60	48.05	39.33	53.85
Median	34.60	34.60	32.10	24.60	38.35
Arbitrage Size/Price					
Mean	0.26	0.28	0.25	0.21	0.27
Median	0.18	0.20	0.17	0.14	0.19
Nadex Trades Considered	1,719	1,413	1,620	954	1,042

Table 6: Silver Arbitrage Results, by Distance to Far CME Strike Price

	Strike Price Distance (dollars/troy ounce)	
	0.05	0.06
Arbitrage Rate	0.32	0.20
Arbitrage Size		
Mean	24.07	35.26
Median	17.10	27.60
Arbitrage Size/Price		
Mean	0.37	0.41
Median	0.29	0.34
Nadex Trades Considered	750	459

strike price distance moves from $0.05 to $0.06.

In conclusion, arbitrage opportunities frequently arise between the Nadex and CME exchanges, with dominating portfolios of CME call options often costing less than their corresponding Nadex binary options. The results hold when we make robustness adjustments to our trade matching algorithm, and arbitrage rates fall as strike price distances increase and arbitrage opportunities become more severe.

6 Evaluating Standard Explanations

The arbitrage opportunities we document could arise either from institutional trading frictions between the Nadex and CME exchanges or from behavioral biases among binary option traders. Among the institutional factors that could account for arbitrage in our setting, we view the most plausible explanations as random price noise, differential trading costs, and differential investor knowledge. If volatile trading among relatively inexperienced retail investors on the relatively thin Nadex exchange causes large, random price fluctuations, then the prices of Nadex options may temporarily move outside of no-arbitrage bounds even if Nadex traders are not acting behaviorally. Similarly, even though we have already accounted for explicit trading fees, if implicit costs make trading at CME more costly than trading at Nadex, then rational retail investors may pass up apparent arbitrage opportunities at CME rather than incur these higher trading costs. And finally, Nadex's marketing of binary options may attract retail investors who are not aware of the traditional options traded at CME and are thus unable to act on arbitrage opportunities, even if they do not exhibit behavioral biases when valuing option payoffs.

After rejecting random price noise, trading costs, and differential knowledge as complete explanations and concluding that behavioral biases must play a role, we show that canonical decision theories like prospect theory and cumulative prospect theory cannot predict our results.

6.1 Random Price Volatility

We first consider the random price noise explanation. If particular market events or short periods of low liquidity on the Nadex exchange caused price volatility that in turn temporarily moved prices outside of no-arbitrage bounds, then we might expect to see spikes in arbitrage rates during particular parts of our sample period. Similarly, if market thinness during particular parts of the trading day caused excess price volatility, we would expect to see increases in arbitrage rates during certain trading hours.

Figures 4-7 investigate these possibilities. In Figures 4-6, we plot the daily arbitrage rate (which, as defined in Section 5, is the fraction of Nadex trades that cost more than their dominating CME portfolio), along with a smoothed local linear trend, for all three option types. Gold and silver arbitrage rates are slightly higher early in the sample period and S&P arbitrage rates are slightly higher during the middle of the sample period, but these time trends are mild and are not driven by a small number of days with outlying arbitrage rates.

Figure 4: S&P Arbitrage Rates by Date

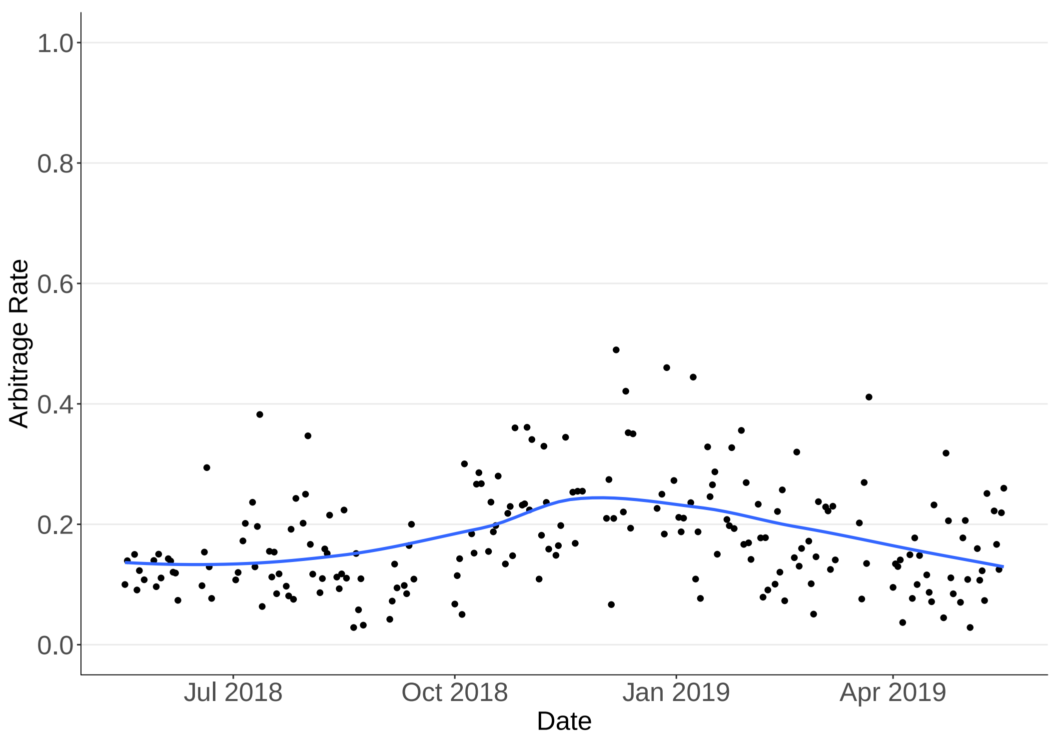


FIGURE 5: Gold Arbitrage Rates by Date

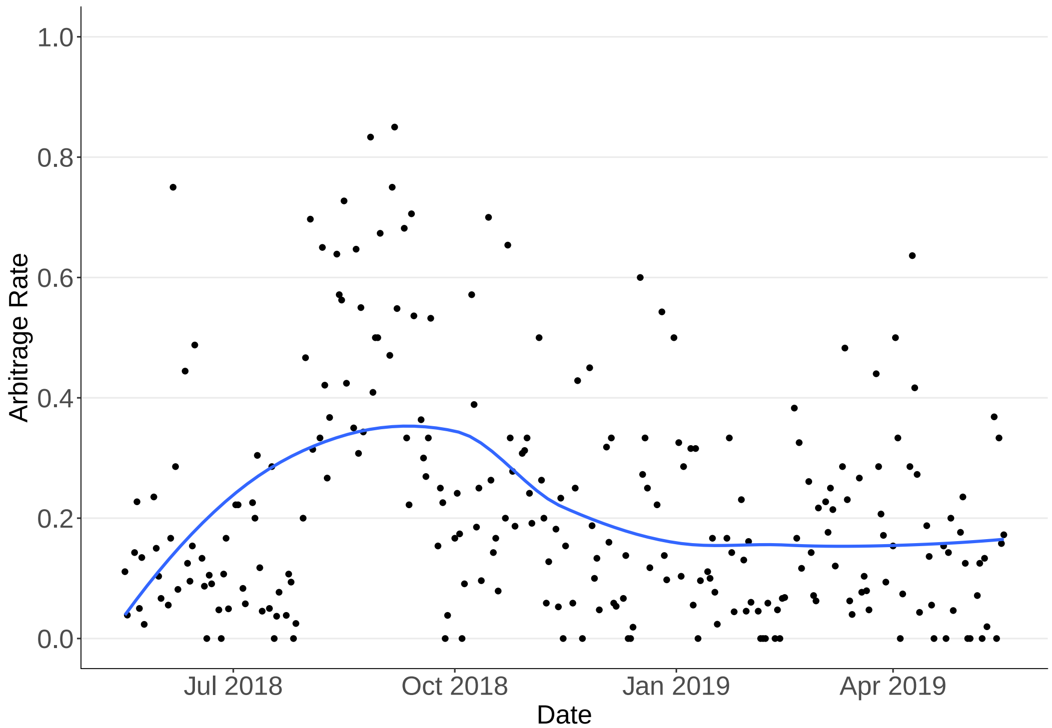


FIGURE 6: Silver Arbitrage Rates by Date

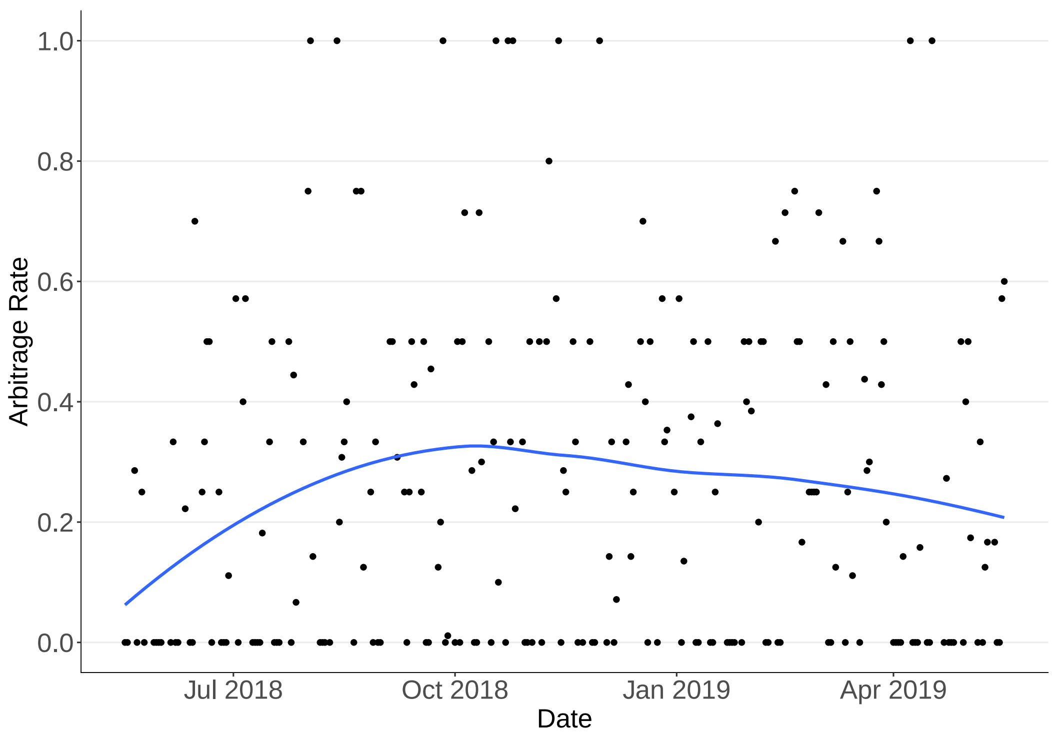
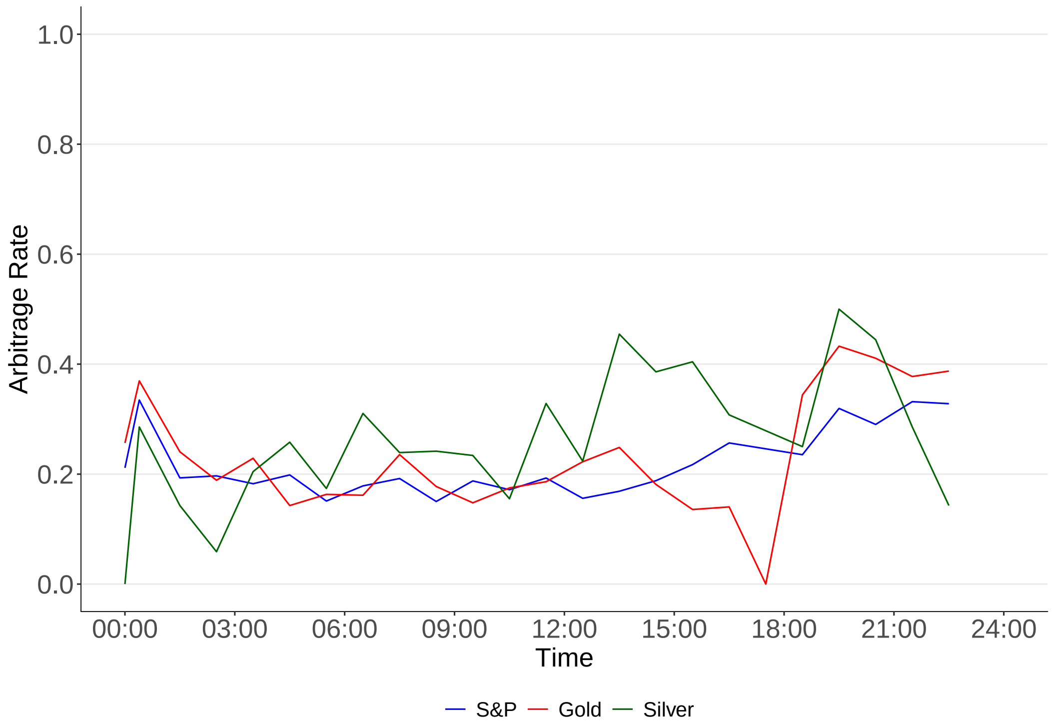


Figure 7: Arbitrage Rates by Time of Day



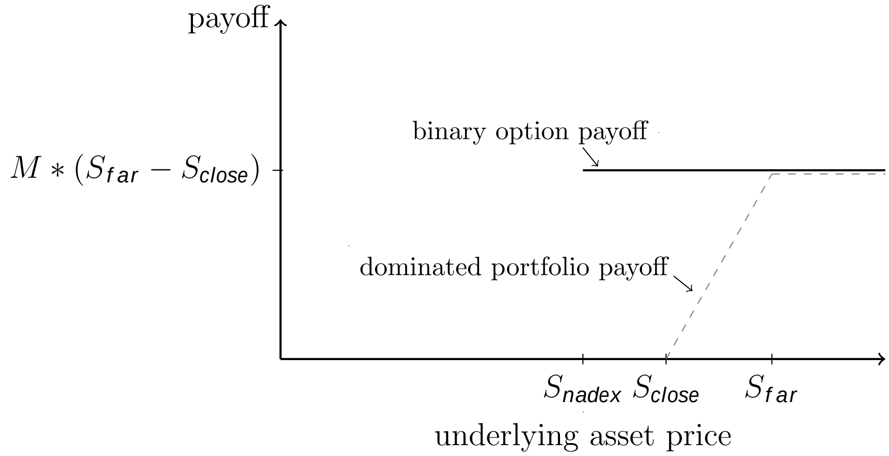
Daily silver arbitrage rates do reach up to 100%, but this occurs on many days that are roughly uniformly distributed over the sample period.

Figure 7 plots average hourly arbitrage rates (across all trading days in the sample period). For all three option types, arbitrage rates are roughly flat during standard trading hours but show a noticeable jump around 18:00; this may relate to the fact that Nadex trading resumes at 18:00 after a short break in the late afternoon. There is thus some evidence that arbitrage rates may be driven partly by market thinness or price volatility during after-hours trading, though the effect is relatively modest and is hard to ascribe directly to random price volatility on Nadex (for example, it may just be that retail investors with behavioral biases are more likely to trade after hours).

We can also address the random price volatility explanation with a kind of placebo test. If the arbitrage rates we observe result only from random price volatility, then arbitrage opportunities should exist just as often on the Nadex exchange as they do on the CME exchange. In other words, our empirical strategy thus far has constructed dominating CME portfolios and looked for cases where they cost less than their corresponding Nadex options, but we could also reverse this logic by constructing ***dominated*** CME portfolios and looking for cases where they cost ***more*** than their corresponding Nadex options. Random price noise should cause both types of arbitrage opportunities to arise at roughly the same rate.

Figure 8 illustrates our placebo test. If a Nadex binary option has strike price S_{nadex},

Figure 8: A Dominated CME Portfolio



we can construct a ***dominated*** CME portfolio by buying a CME call option at S_{close} and selling a CME call option at S_{far}. For each Nadex trade in our sample, we re-run our matching algorithm to look for CME price quotes at strike prices just above the Nadex strike price, construct a dominated CME portfolio as shown in Figure 8, and determine how often the dominated CME portfolio costs more. The results are shown in Table 7: columns 1-3 reproduce our baseline results where the CME portfolio is dominating, while columns 4-6 report the new placebo results where the CME portfolio is dominated. Though arbitrage rates are still nonzero in the placebo test, they are substantially smaller than the baseline results: the S&P arbitrage rate decreases from 19% to 14%, the gold arbitrage rate from 21% to 7%, and the silver arbitrage rate from 25% to 17%. As indicated by the reported t-statistics, the decrease in the arbitrage rate in the placebo test is highly significant (at the 1% level) for all three asset classes. The placebo results indicate that price noise is responsible for some, but not all, of our baseline arbitrage rates, and point toward a behavioral explanation that justifies a strict preference for binary options over comparable call option portfolios.

6.2 Trading Costs

We also consider the possibility of differential trading costs between the exchanges. As discussed above, we have already accounted for explicit trading fees when identifying cases of arbitrage. In derivatives markets, collateral requirements are another important source of trading costs, as exchanges often require traders to post collateral (in proportion to the current value or riskiness of their positions) in order to minimize the risk of counterparty default. But note that since both the payoff profiles of a binary option and its dominating

Table 7: Placebo Test with Dominated CME Portfolios

	CME is Dominating			CME is Dominated		
	S&P (1)	Gold (2)	Silver (3)	S&P (4)	Gold (5)	Silver (6)
Arbitrage Rate	0.19	0.21	0.25	0.14	0.07	0.17
t-Statistic for Difference				(18.79)	(23.31)	(5.30)
Arbitrage Size						
Mean	41.28	51.01	28.30	40.80	50.45	40.00
Median	21.90	33.35	19.60	26.10	37.90	25.90
Arbitrage Size/Price						
Mean	0.34	0.26	0.38	0.25	0.24	0.29
Median	0.29	0.18	0.30	0.18	0.15	0.23
Nadex Trades Considered	33,676	6,876	1,355	34,552	6,791	1,522

Note: Only Nadex trades that were successfully matched to a CME portfolio are considered. Columns 1-3 present arbitrage measures when the CME portfolio is constructed to dominate the Nadex option (the same baseline results shown in Table 2); arbitrage in these cases occurs when the CME portfolio costs less. Columns 4-6 present arbitrage measures when the CME portfolio is constructed to be dominated by the Nadex option; arbitrage in these cases occurs when the CME portfolio costs more. See notes to Table 3 for definitions of the arbitrage measures. The t-statistic reported in the second row tests the hypothesis that the arbitrage rate for each asset class is the same for the dominating-CME and dominated-CME cases.

CME portfolio are always weakly positive (see Figure 2), in neither case can the holder of the position lose more money than the initial trade price. As a result, neither Nadex nor CME requires traders to post collateral when maintaining the positions we consider in this paper.[10]

However, for completeness, we acknowledge the possibility that the two trades necessary to construct the CME portfolio may be executed sequentially and temporarily create loss exposure for the trader, and that the brokers who give retail traders access to CME may set their own collateral requirements. To ensure the robustness of our findings to any possible difference in collateral requirements between the exchanges, in Table 8 we determine whether measures of price volatility on CME predict arbitrage rates. During periods of high volatility, CME's collateral-requirement algorithm will perceive higher risk and demand more collateral from traders. This makes trading at CME more costly relative to Nadex, and if differential collateral requirements are actually driving our results, we would expect to see arbitrage rates increase as a result (since higher CME trading costs would deter traders from acting

[10] See official exchange material on collateral requirements at Nadex and CME.

TABLE 8: Arbitrage Rates and CME Market Volatility

	Panel A: S&P							
	Same-Day				Lagged			
	Max-Min (1)	P95-P5 (2)	Variance (3)	Hourly Diff. (4)	Max-Min (5)	P95-P5 (6)	Variance (7)	Hourly Diff. (8)
Estimate	0.002	0.005	0.001	0.019	0.002	0.004	0.001	0.197
	(2.628)	(6.216)	(3.953)	(2.045)	(4.720)	(3.207)	(2.316)	(2.136)
R-squared	0.138	0.195	0.129	0.017	0.135	0.082	0.054	0.044
Marginal Effect	0.033	0.039	0.032	0.012	0.033	0.025	0.021	0.019
	Panel B: Gold							
	Same-Day				Lagged			
	Max-Min	P95-P5	Variance	Hourly Diff.	Max-Min	P95-P5	Variance	Hourly Diff.
Estimate	0.004	0.013	0.005	-0.033	0.000	0.003	-0.000	-0.608
	(1.568)	(2.563)	(1.880)	(-3.857)	(0.083)	(0.593)	(-0.020)	(-1.640)
R-squared	0.009	0.023	0.010	0.011	0.000	0.001	0.000	0.007
Marginal Effect	0.018	0.029	0.019	-0.020	0.001	0.007	-0.000	-0.016
	Panel C: Silver							
	Same-Day				Lagged			
	Max-Min	P95-P5	Variance	Hourly Diff.	Max-Min	P95-P5	Variance	Hourly Diff.
Estimate	0.055	-0.177	3.497	0.074	0.177	-0.419	-2.848	0.209
	(0.841)	(-0.346)	(0.494)	(0.406)	(2.403)	(-0.992)	(-0.880)	(0.794)
R-squared	0.003	0.001	0.002	0.001	0.033	0.003	0.001	0.001
Marginal Effect	0.016	-0.007	0.012	0.006	0.051	-0.015	-0.009	0.010

Note: Only Nadex trades that were successfully matched to a CME portfolio are considered. The table presents results from simple bivariate regressions where the observations are contract-days, the dependent variable is the contract-specific daily arbitrage rate, and the independent variable is a measure of price volatility on the CME exchange. The max-min spread is the difference between the day's highest and lowest observed prices, the p95-p5 spread is the difference between the 95th and 5th percentiles of the day's observed prices, variance is computed over all of the day's observed prices, and hourly diff. gives the mean difference between consecutive hourly average prices during the day. Columns 1-4 use CME volatility measures from the same day on which the arbitrage rate is computed, while columns 5-8 use CME volatility measures from the previous trading day. In addition to coefficient estimates, t-statistics (in parentheses), and R-squared values, the table shows the predicted marginal effect on the daily arbitrage rate of a one-standard deviation increase in the independent variable. Standard errors are clustered by contract, where contracts are defined by their underlying asset (S&P, gold, or silver) and expiration date. CME volatility measures are averaged over all listed strike prices for a contract.

on arbitrage opportunities).

Table 8, which examines the relationship between arbitrage rates and several different CME volatility measures in simple bivariate models, gives some mixed results. While there is little evidence that any of the volatility measures are significantly positively correlated with gold or silver arbitrage rates, there is a clear positive relationship between CME market volatility and S&P arbitrage rates. However, despite the positive coefficient estimates, the R-squared values in the S&P panel range between 0.017 and 0.195, indicating that most of the variation in daily S&P arbitrage rates remains unexplained. And once again, given that the payoffs of our CME portfolios are always weakly positive, we do not think that implicit trading costs from collateral requirements are actually important in this setting. We conduct the tests in Table 8 out of an abundance of caution, and note that CME market volatility may itself directly increase arbitrage rates if Nadex prices are slow to respond to CME price movements.

Market liquidity may be another factor that creates differences in implicit trading costs between the two exchanges. Trading in a thin market is costly, since traders who wish to exit their positions before expiration face a larger bid-ask spread and must pay a larger implicit cost by accepting a less favorable liquidation price. Investors should prefer to trade in more liquid markets, and should be willing to pay a premium (in the form of higher asset prices) in order to do so. While we are not able to compare Nadex and CME bid-ask spreads directly,[11] CME is very clearly the more liquid exchange: many more market makers compete to set prices, many more institutional investors have access, and trading volume is orders of magnitude higher than at Nadex. CME's superior liquidity therefore makes our arbitrage results even more surprising: investors should be willing to pay a premium in order to trade there, but we find that dominating CME portfolios often cost less than Nadex binary options. In other words, liquidity considerations actually work against us and make the arbitrage opportunities we detect less likely to occur.

While we can reject differential collateral requirements and market liquidity as explanations for our results, Nadex and CME do differ in terms of their direct accessibility to retail traders. Investors can open an account and trade directly on Nadex's website, but they can only access CME by opening an account with a third-party broker. These brokers may charge trading fees or other types of transaction costs in excess of CME's own trading fees, thereby increasing the cost of trading on CME relative to Nadex. Gauging the extent of additional trading fees in the decentralized market for CME brokers is difficult, but we doubt that they are large enough to erase the sizable arbitrage opportunities we identify here (recall from Table 2 that the mean arbitrage sizes for S&P, gold, and silver options are \$41.28, \$51.01, and \$28.30, respectively).

6.3 Nadex Marketing

Given that Nadex actively attempts to attract small retail investors to its website,[12] it is possible that the company's marketing efforts help create our observed arbitrage opportunities by asymmetrically informing traders about binary options. In this section, we conduct a price-based test of this hypothesis. Later, in Section 8, we directly ask binary options traders about their knowledge of and ability to trade traditional option contracts.

A simple story would be the following: retail investors who were not previously trading any types of options discover, through one of Nadex's online advertisements, that novel

[11] Our Nadex data do not indicate the side of the market on which trades took place, so we cannot infer bid-ask spreads.

[12] See the kind of marketing language employed on Nadex's landing page for binary option trading, which the company also uses in search advertising.

Table 9: Arbitrage Rates and Google Search Activity for ‘Nadex’

	Panel A: S&P				
	Lag Period				
	1	2	3	4	5
Lagged Search Activity	-0.000	0.000	-0.000	-0.001	-0.002
	(-0.236)	(0.371)	(-0.563)	(-0.980)	(-1.942)
	Panel B: Gold				
	Lag Period				
	1	2	3	4	5
Lagged Search Activity	-0.000	-0.001	-0.001	-0.003	-0.004
	(-0.040)	(-0.680)	(-1.119)	(-1.950)	(-2.232)
	Panel C: Silver				
	Lag Period				
	1	2	3	4	5
Lagged Search Activity	-0.001	-0.003	-0.003	-0.001	-0.000
	(-0.683)	(-2.068)	(-1.493)	(-0.553)	(-0.014)

Note: Only Nadex trades that were successfully matched to a CME portfolio are considered. The table presents results from simple bivariate regressions where the dependent variable is the contract-specific daily arbitrage rate, and the independent variable is the lagged average level of daily Google search activity for the word “Nadex” (reported on a scale of 1 to 100). Lag periods of 1 to 5 days are considered, with a lag period of 1 corresponding to the contemporaneous day on which the arbitrage rate is measured. T-statistics are in parentheses, and standard errors are clustered by contract, where contracts are defined by their underlying asset (S&P, gold, or silver silver futures) and expiration date.

binary option contracts are available to be traded at Nadex. To locate the Nadex website or learn more about the exchange, they enter “Nadex” into a search engine. They then open a Nadex trading account and begin trading binary options within the next few days. As a result, there are now more retail traders at Nadex who are unaware of traditional option contracts and are not able to compare price quotes between the Nadex and CME platforms. The entry of these new traders to the Nadex platform causes arbitrage rates to increase.

Under the story sketched above, we would expect Nadex-related Google searches to correlate positively with subsequent arbitrage rates. Our test of the marketing explanation is thus to regress arbitrage rates on measures of Google search activity for the word “Nadex.” The results are presented in Table 9. The dependent variable is the contract-specific daily arbitrage rate, and the independent variable is the average level of Google search activity over lagged periods of 1 to 5 days (with day 1 defined as the contemporaneous day on which the arbitrage rate is measured). Google reports daily search activity on a scale of 1 to 100, with 100 corresponding to the highest number of daily searches during our sample period. The point estimates in Table 9 are almost all negative (and the single positive estimate is insignificant), indicating that higher Google search activity does not correlate with higher arbitrage rates. This test thus does not yield any evidence that our arbitrage results are caused by marketing-driven information asymmetries among inexperienced binary option traders.

6.4 Canonical Decision Theories

If our arbitrage results are not driven by random price volatility, trading costs, targeted Nadex marketing, or any other institutional factor, the explanation may lie in some sort of departure of retail traders' preferences from standard expected utility theory. But cumulative prospect theory, and most other well-known revisions to the expected utility model, do not allow for violations of dominance (Kahneman and Tversky, 1992). The arbitrage result in this case is a violation of dominance.

While prospect theory (Kahneman and Tversky, 1979) allows for violations of dominance, it does not accord with our empirical findings; prospect theory agents would never choose a dominated binary option over a dominant call option portfolio.

To see that prospect theory cannot explain our results, observe that for any probability weighting function π and reference point R, the utility a prospect theory agent receives from a portfolio of binary options or call options is

$$U = \int \pi(p(s))u_R(q(s))ds, \tag{7}$$

where s is the underlying asset price at contract expiration, u_R is a utility function with the R subscript denoting dependence on the reference point, $p(s)$ is the probability of underlying asset price s occurring, and $q(s)$ is the monetary payoff that the agent receives from the option portfolio when underlying asset price s occurs. Consider a prospect theory agent who buys a binary option with strike price K. Since the binary option pays \$100 if $s > K$ and zero otherwise, the utility of the binary option is

$$U(B) = \int_0^K \pi(p(s))u_R(0)ds + \int_K^1 \pi(p(s))u_R(100)ds. \tag{8}$$

In contrast, the comparable call option porfolio O in our empirical test (constructed by buying a call option with strike $K - \epsilon$ and selling a call option with strike K) gives utility[13]

$$U(O) = \int_0^K \pi(p(s))u_R(0)ds + \int_K^K \pi(p(s))u_R(s - (K - \epsilon))ds + \int_K^1 \pi(p(s))u_R(100)ds. \tag{9}$$

Because $\pi(\cdot)$ is weakly positive and $u_R(\cdot)$ weakly increasing in prospect theory, it follows

[13]For simplicity but without loss of generality, we assume that the notional value of the call options is such that the portfolio pays exactly 100 if the underlying asset price is above K at expiration.

that, for any $\epsilon > 0$,

$$\begin{aligned}\int_0^K \pi(p(s))u_R(0)ds + \int_K^K \pi(p(s))u_R(s-(K-\epsilon))ds \\ \geq \int_0^K \pi(p(s))u_R(0)ds + \int_K^K \pi(p(s))u_R(0)ds \\ = \int_0^K \pi(p(s))u_R(0)ds, \end{aligned} \tag{10}$$

and therefore that $U(O) \geq U(B)$ for any prospect theory agent, regardless of functional form. This shows that a prospect theory agent cannot strictly prefer a dominated binary option to a comparable call option portfolio, and thus that prospect theory cannot explain our arbitrage results. In the next section, we discuss a novel decision theory that does predict our empirical results.

7 Alternative Explanation: Preference for Simplicity

Simplicity theory, introduced in Puri (2020), is a model of choice under risk in which the size of a lottery's support, as well as its expected utility, matters in agents' preference rankings. We summarize the theory here and show that, unlike the standard explanations considered in Section 6, it is capable of predicting the observed arbitrage results. We note that while we use the term 'preference for simplicity,' we, like the paper introducing the theory, are agnostic as to whether this behavior constitutes a preference or a mistake. The word preference here refers only to a revealed preference.

Let X be the relevant outcome set and $x \in X$ be elements of the outcome set. Further, let $\Delta(X)$ be the set of lotteries over X that have finite support. In our application X is the set of possible monetary payoffs provided by binary options or portfolios of traditional call options. Note that simplicity preferences are only defined over lotteries with finite support, and in theory the set of possible call option payoffs is a continuum. However, in practice the underlying assets of the options we study (S&P index, gold, and silver futures) have discrete minimum price fluctuations, which ensures that all lotteries we consider have finite support.[14] Finally, let $p \in \Delta(X)$ be lotteries and $u(\cdot)$ be the Bernoulli utility function.

[14] The underlying assets for both the binary options and traditional call options we study are futures contracts traded at CME. According to CME trading rules, the minimum price fluctuations for S&P, gold, and silver futures contracts are 0.25 index points, \$0.10 per troy ounce, and \$0.005 per troy ounce, respectively. The payoffs of both binary options and call options are bounded below by zero, and in our application the set of possible payoffs is also bounded above (see Figure 2). Thus all possible outcomes of the lotteries we consider belong to a bounded, discrete set, which in turn means that all lotteries in $\Delta(X)$ have finite support.

The preferences of an agent who obeys the axioms of simplicity theory can be represented by

$$U(p) = \sum p(x)u(x) - C(|\text{support}(p)|), \tag{11}$$

where $U(p)$ is the utility of lottery p and $C : \mathbb{Z}^+ \mapsto \mathbb{R}$ is a weakly positive, weakly increasing function. Intuitively, a simplicity agent has standard expected utility preferences but also penalizes lotteries with larger supports through the cost function C.

The simplicity representation of preferences leads to three key theoretical predictions that accord with our empirical analysis. First and most clearly, because of the increasing cost function C, simplicity agents assign a utility premium to lotteries with fewer possible outcomes. When given a choice between two lotteries that provide equal expected payoff utility, a simplicity agent strictly prefers the one with a smaller support.

Second, simplicity theory rationalizes choices that expected utility theory, prospect theory, and cumulative prospect theory cannot. In particular, a simplicity agent may violate dominance, strictly preferring a small-support lottery q to a large-support lottery p even when lottery p provides a weakly higher payoff in every state of the world. Expected utility theory and cumulative prospect theory do not allow for violations of dominance, and while dominance violations can occur under prospect theory, in our setup a prospect theory agent would never strictly prefer a binary option to a dominating CME portfolio (see Section 6.4).

Third, simplicity agents are more likely to violate dominance when the amount of dominance is "small." More concretely, suppose that lottery p dominates lottery q, but that $\sum p(x)u(x)$ is only slightly larger than $\sum q(x)u(x)$. This case could arise when lotteries p and q are largely the same, except for a small number of possible events where p gives a higher payoff than q. Since the expected utility components of p and q are very similar, the cost-function difference $C(\text{support}(p)) - C(\text{support}(q))$ will be more important in determining which lottery the simplicity agent prefers. If p has even a slightly larger support than q, this complexity cost may be enough to outweigh p's expected utility advantage and lead a simplicity agent to prefer q. In other words, differences in lottery supports become more consequential, and more likely to lead to dominance violations, when the available lotteries' expected utility components are more similar.

Our empirical results directly accord with the hypotheses of simplicity theory. Table 2 establishes the main result that the prices of binary options are often higher than the prices of dominating portfolios of call options. Given that we can reject price volatility (Section 6.1), trading costs (Section 6.2), and prospect theory (Section 6.4) as full explanations, we are led to conclude that at least some retail traders prefer smaller lottery supports (the first prediction of simplicity theory) and are willing to violate dominance in order to purchase contracts with fewer possible outcomes (the second prediction of simplicity theory).

Regarding the final prediction, we show that arbitrage rates consistently decrease as the distance between the Nadex strike price and the far strike price in the CME portfolio (see again Figure 2) increases. This result aligns closely with the above discussion. When the strike price distance is smaller, the CME portfolio provides a strictly higher payoff over a smaller range of underlying asset prices and is thus "less dominant," providing an expected utility component more similar to that of the binary option. This means that the complexity cost is more consequential in determining preference rankings, and that dominance violations are more likely, when the strike price distance is small. Simplicity theory is thus consistent with the clear negative relationship between strike price distance and arbitrage rates that we document in Tables 4-6.

8 Direct Survey Evidence from Binary Option Traders

With market trading data and straightforward theoretical reasoning, we have been able to reject standard institutional factors and canonical decision theories as complete explanations for our arbitrage results, and to establish simplicity theory as an alternative behavioral model capable of generating the dominance violations we observe. We now strengthen our conclusions with two pieces of direct evidence that can only be obtained from a survey of binary option traders.

First, we corroborate our interpretation of the trading data with controlled survey questions that test directly for the presence of simplicity preferences among binary option traders. The discussion in Section 7 shows that our empirical arbitrage results are consistent with the key predictions of simplicity theory, but is not an explicit proof that Nadex traders do indeed possess simplicity preferences. Surveys have lower stakes and yield smaller sample sizes than real-world trading environments, but allow a direct test of the simplicity-theory mechanism we propose to explain our observed dominance violations.

Second, we verify that binary option traders have sufficient institutional knowledge to identify the arbitrage opportunities we document. If Nadex traders are simply not aware that traditional call options are available at CME, then they would not be able to act on these arbitrage opportunities and our findings could be driven by the resulting segmentation of the Nadex and CME markets rather than by simplicity preferences. Our survey allows us to pose direct questions confirming that Nadex traders are able to purchase CME call option portfolios and choose not to.

8.1 Methodology

We posted our survey on the popular binary option trading forum trade2win.com during the week of August 17, 2020. Respondents were offered a $10 base payment for completing the survey, as well as a bonus payment resulting from the simulation of a randomly chosen lottery-preference question, which served to incentivize truthful preference reporting. The survey includes five lottery-choice questions (referred to as Questions 1-5 and shown in Figures 9-11). These five questions include two tests for simplicity theory, two tests for cumulative prospect theory, and one test for prospect theory. The tests are summarized in Table 10 and explained in more detail below. Participants received standard instructions for the lottery-choice questions (shown in Appendix B.2) and question order was randomized.[15]

In addition to the lottery-choice questions used to elicit preferences, we included questions about respondents' basic demographic information, experience trading binary options, and knowledge of and access to the traditional option contracts traded at CME. The forum post we used to advertise the survey, as well as the full set of survey questions, are given in Appendix B.1.

Following standard procedure, to ensure clean data, we drop subjects who display any instance of multiple switching in their responses to the lottery-choice questions. As a final attention check, because the survey was posted on a binary option trading forum, we drop any individuals who indicate that they have never heard of binary options. This yields a final sample size of 118 respondents.

8.2 Preference Tests

In this section we describe the design of our preference-elicitation tests. We include questions that directly test the predictions of simplicity theory, prospect theory, and cumulative prospect theory. Additionally, because one of our tests gives respondents the option of choosing a strictly dominated lottery, it is capable of rejecting any theory that respects dominance. Because we analyze differences in contract pricing that may arise from the smaller supports of binary option lotteries, we focus on tests that ask respondents to express preferences over lotteries with different numbers of outcomes.

8.2.1 Event Splitting

Event splitting refers, in this case, to the agent's preference between lotteries ($2, 60%; $3, 40%) and ($1.50, 30%; $2.50, 30%; $3, 40%). Prospect theory and simplicity theory make

[15] Precisely, the order was either: 1,2,3,5,4 or 4,5,3,2,1, and the order was chosen uniformly at random.

Table 10: Preference Tests

	Test 1	Test 2
Simplicity	Event splitting in favor of fewer outcomes	Violates dominance in favor of fewer outcomes
CPT	Displays rank dependence	Respects dominance
PT	Event splitting in favor of more outcomes	–

opposite predictions. A simplicity agent assigns a weakly positive utility premium to fewer outcomes, and therefore weakly prefers the first lottery. Under prospect theory, the S-shaped weighting function implies that more weight is given to the 60% outcome than to the two 30% outcomes combined. A PT agent therefore prefers the second lottery. Earlier literature using similar questions finds evidence in favor of the simplicity prediction (Bernheim and Sprenger, 2020).

Question 1 elicits a certainty equivalent for the first lottery, and Question 2 elicits a certainty equivalent for the second lottery. The first test for simplicity theory is whether the certainty equivalent for the first lottery is weakly higher than the certainty equivalent for the second. The first test for prospect theory is whether the opposite phenomenon occurs.

8.2.2 Rank Dependence

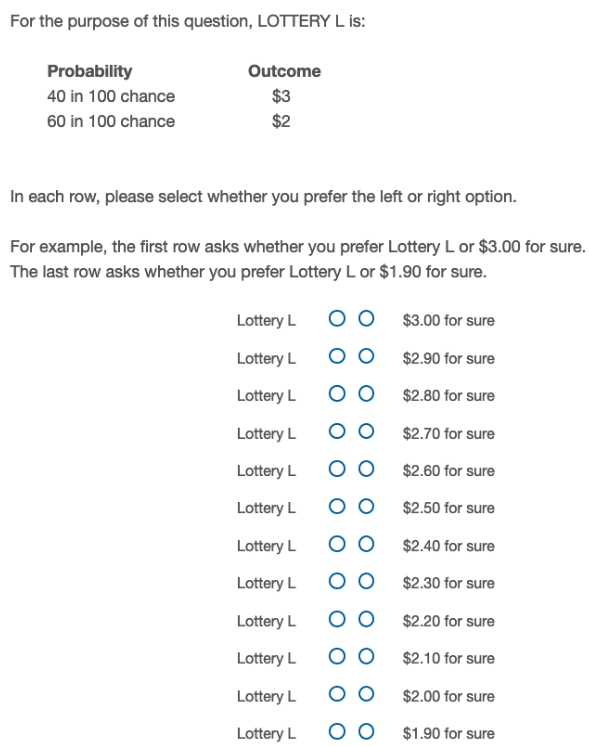
Cumulative prospect theory assigns different probability weighting functions to the best and second-best outcomes. A standard test for CPT is therefore whether this rank dependent probability weighting holds. Each of Questions 3 and 4 elicits an equalizing payoff, that is, r_1 and r_2 satisfying the following indifference conditions:

$$\text{Q3: } (\$8.50, 40\%; \$9, 30\%; \$1.50, 30\%) \sim (\$8.50, 40\%; \$9.50, 30\%; \$r_1, 30\%)$$

$$\text{Q4: } (\$10, 40\%; \$9, 30\%; \$1.50, 30\%) \sim (\$10, 40\%; \$9.50, 30\%; \$r_2, 30\%)$$

In words, what is the amount of money that would make the subject indifferent to increasing outcome \$9 to \$9.50?

The rank dependence feature of CPT implies that this amount of money should be different in each question (e.g. $r_1 \neq r_2$). This is because, in Q3, \$9 is originally the best



For the purpose of this question, LOTTERY L is:

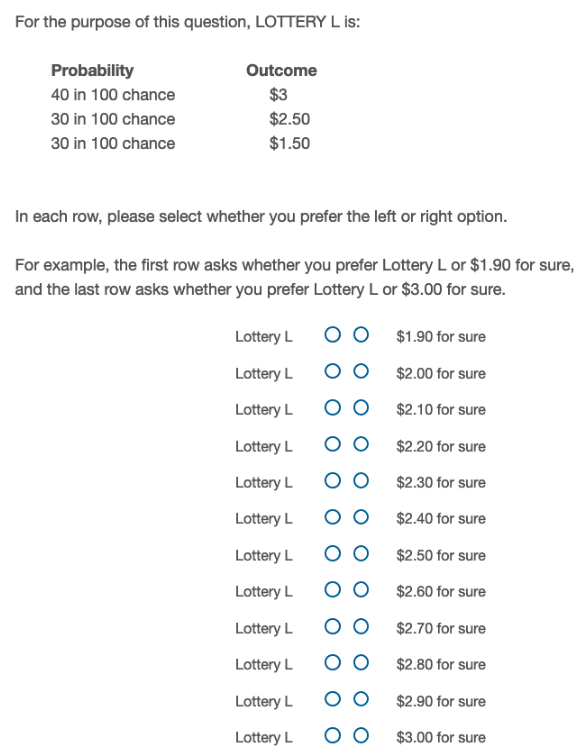
Probability	Outcome
40 in 100 chance	$3
60 in 100 chance	$2

In each row, please select whether you prefer the left or right option.

For example, the first row asks whether you prefer Lottery L or $3.00 for sure. The last row asks whether you prefer Lottery L or $1.90 for sure.

Lottery L	$3.00 for sure
Lottery L	$2.90 for sure
Lottery L	$2.80 for sure
Lottery L	$2.70 for sure
Lottery L	$2.60 for sure
Lottery L	$2.50 for sure
Lottery L	$2.40 for sure
Lottery L	$2.30 for sure
Lottery L	$2.20 for sure
Lottery L	$2.10 for sure
Lottery L	$2.00 for sure
Lottery L	$1.90 for sure

(A)



For the purpose of this question, LOTTERY L is:

Probability	Outcome
40 in 100 chance	$3
30 in 100 chance	$2.50
30 in 100 chance	$1.50

In each row, please select whether you prefer the left or right option.

For example, the first row asks whether you prefer Lottery L or $1.90 for sure, and the last row asks whether you prefer Lottery L or $3.00 for sure.

Lottery L	$1.90 for sure
Lottery L	$2.00 for sure
Lottery L	$2.10 for sure
Lottery L	$2.20 for sure
Lottery L	$2.30 for sure
Lottery L	$2.40 for sure
Lottery L	$2.50 for sure
Lottery L	$2.60 for sure
Lottery L	$2.70 for sure
Lottery L	$2.80 for sure
Lottery L	$2.90 for sure
Lottery L	$3.00 for sure

(B)

Figure 9: Event Splitting Questions

Panel (A) shows Question 1. Panel (B) shows Question 2. While a simplicity agent prefers the lottery in Q1 to that in Q2, a PT agent has the opposite preference.

For the purpose of this question, "LOTTERY L" is:

Probability	Outcome
40 in 100 chance	$8.50
30 in 100 chance	$9
30 in 100 chance	$1.50

And "LOTTERY Q" is:

Probability	Outcome
40 in 100 chance	$8.50
30 in 100 chance	$9.50
30 in 100 chance	$*x*

In each row, please select whether you prefer the left or right option.

For example, the first row asks whether you prefer Lottery L, or Lottery Q with x = $1.50, and the last row asks whether you prefer Lottery L, or Lottery Q with x = $0.

Lottery L — Lottery Q with x = $1.50
Lottery L — Lottery Q with x = $1.40
Lottery L — Lottery Q with x = $1.30
Lottery L — Lottery Q with x = $1.20
Lottery L — Lottery Q with x = $1.10
Lottery L — Lottery Q with x = $1.00
Lottery L — Lottery Q with x = $0.90
Lottery L — Lottery Q with x = $0.80
Lottery L — Lottery Q with x = $0.70
Lottery L — Lottery Q with x = $0.60
Lottery L — Lottery Q with x = $0.50
Lottery L — Lottery Q with x = $0.40
Lottery L — Lottery Q with x = $0.30
Lottery L — Lottery Q with x = $0.20
Lottery L — Lottery Q with x = $0.10
Lottery L — Lottery Q with x = $0.00

(A)

For the purpose of this question, "LOTTERY L" is:

Probability	Outcome
40 in 100 chance	$10
30 in 100 chance	$9
30 in 100 chance	$1.50

And "LOTTERY Q" is:

Probability	Outcome
40 in 100 chance	$10
30 in 100 chance	$9.50
30 in 100 chance	$*x*

In each row, please select whether you prefer the left or right option.

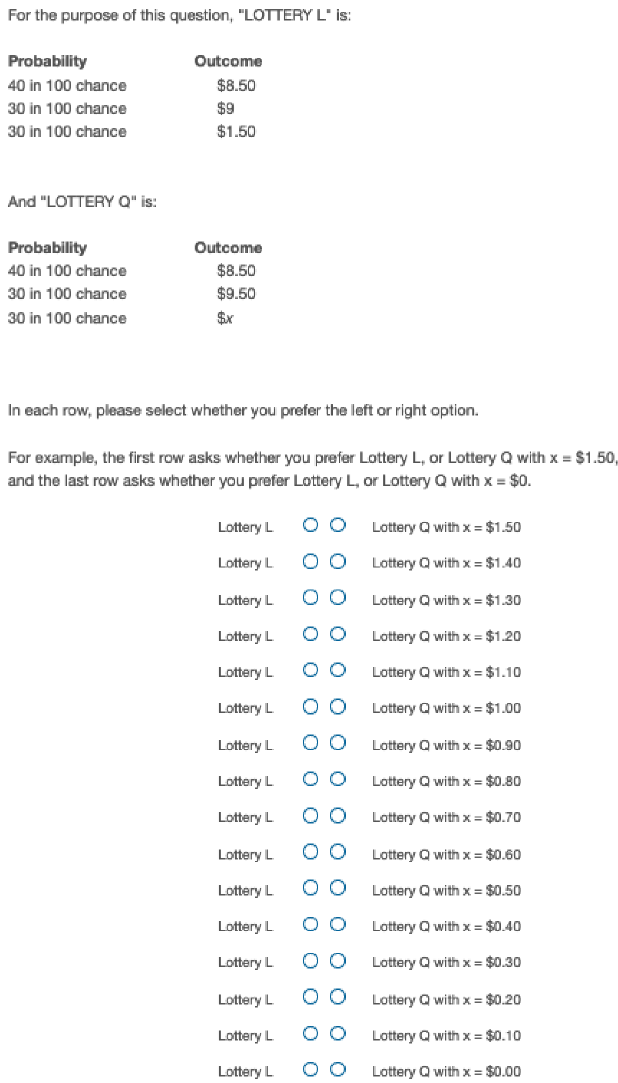
For example, the first row asks whether you prefer Lottery L, or Lottery Q with x = $1.50, and the last row asks whether you prefer Lottery L, or Lottery Q with x = $0.

Lottery L — Lottery Q with x = $1.50
Lottery L — Lottery Q with x = $1.40
Lottery L — Lottery Q with x = $1.30
Lottery L — Lottery Q with x = $1.20
Lottery L — Lottery Q with x = $1.10
Lottery L — Lottery Q with x = $1.00
Lottery L — Lottery Q with x = $0.90
Lottery L — Lottery Q with x = $0.80
Lottery L — Lottery Q with x = $0.70
Lottery L — Lottery Q with x = $0.60
Lottery L — Lottery Q with x = $0.50
Lottery L — Lottery Q with x = $0.40
Lottery L — Lottery Q with x = $0.30
Lottery L — Lottery Q with x = $0.20
Lottery L — Lottery Q with x = $0.10
Lottery L — Lottery Q with x = $0.00

(B)

FIGURE 10: Rank Dependence Questions

Panel (A) shows Question 3. Panel (B) shows Question 4. A CPT agent will choose a higher equalizing $x for Q4 than Q3.



For the purpose of this question, "LOTTERY L" is:

Probability	Outcome
40 in 100 chance	$8.50
30 in 100 chance	$9
30 in 100 chance	$1.50

And "LOTTERY Q" is:

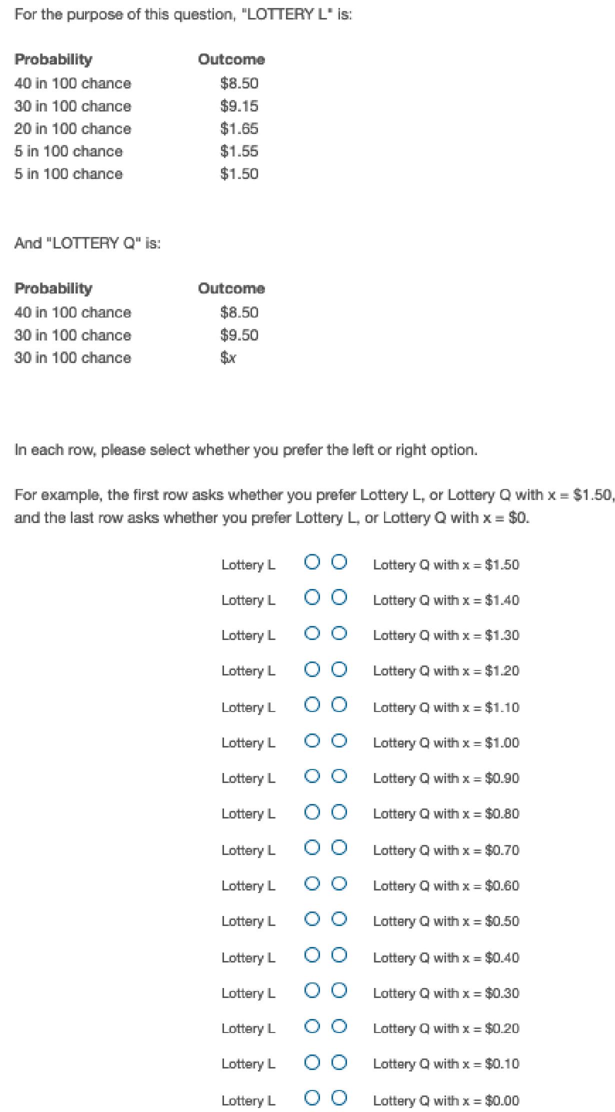
Probability	Outcome
40 in 100 chance	$8.50
30 in 100 chance	$9.50
30 in 100 chance	$*x*

In each row, please select whether you prefer the left or right option.

For example, the first row asks whether you prefer Lottery L, or Lottery Q with x = $1.50, and the last row asks whether you prefer Lottery L, or Lottery Q with x = $0.

Lottery L	Lottery Q with x = $1.50
Lottery L	Lottery Q with x = $1.40
Lottery L	Lottery Q with x = $1.30
Lottery L	Lottery Q with x = $1.20
Lottery L	Lottery Q with x = $1.10
Lottery L	Lottery Q with x = $1.00
Lottery L	Lottery Q with x = $0.90
Lottery L	Lottery Q with x = $0.80
Lottery L	Lottery Q with x = $0.70
Lottery L	Lottery Q with x = $0.60
Lottery L	Lottery Q with x = $0.50
Lottery L	Lottery Q with x = $0.40
Lottery L	Lottery Q with x = $0.30
Lottery L	Lottery Q with x = $0.20
Lottery L	Lottery Q with x = $0.10
Lottery L	Lottery Q with x = $0.00

(A)



For the purpose of this question, "LOTTERY L" is:

Probability	Outcome
40 in 100 chance	$8.50
30 in 100 chance	$9.15
20 in 100 chance	$1.65
5 in 100 chance	$1.55
5 in 100 chance	$1.50

And "LOTTERY Q" is:

Probability	Outcome
40 in 100 chance	$8.50
30 in 100 chance	$9.50
30 in 100 chance	$*x*

In each row, please select whether you prefer the left or right option.

For example, the first row asks whether you prefer Lottery L, or Lottery Q with x = $1.50, and the last row asks whether you prefer Lottery L, or Lottery Q with x = $0.

Lottery L	Lottery Q with x = $1.50
Lottery L	Lottery Q with x = $1.40
Lottery L	Lottery Q with x = $1.30
Lottery L	Lottery Q with x = $1.20
Lottery L	Lottery Q with x = $1.10
Lottery L	Lottery Q with x = $1.00
Lottery L	Lottery Q with x = $0.90
Lottery L	Lottery Q with x = $0.80
Lottery L	Lottery Q with x = $0.70
Lottery L	Lottery Q with x = $0.60
Lottery L	Lottery Q with x = $0.50
Lottery L	Lottery Q with x = $0.40
Lottery L	Lottery Q with x = $0.30
Lottery L	Lottery Q with x = $0.20
Lottery L	Lottery Q with x = $0.10
Lottery L	Lottery Q with x = $0.00

(B)

FIGURE 11: Dominance Violation Questions

Panel (A) shows Question 3. Panel (B) shows Question 5. Lottery L in Q5 dominates Lottery L in Q3. A CPT agent will choose a strictly lower equalizing $x for Q5 than for Q3. A simplicity agent may choose a weakly higher equalizing $x for Q5 than for Q3.

outcome, while in Q4 it is the second-best outcome. In fact, for canonical CPT,[16] $r_1 < r_2$ (Bernheim and Sprenger, 2020). This is the first test for CPT.

8.2.3 Violations of Dominance

Cumulative prospect theory predicts no violations of dominance. A simplicity agent with sufficiently strong complexity aversion may violate dominance in favor of lotteries with fewer outcomes (see the discussion in Section 7). The final pair of questions compares a dominating five-outcome lottery (Question 5) to a dominated three-outcome lottery (Question 3). A CPT agent will prefer the lottery in Q5 to the lottery in Q3, while a simplicity agent may not respect dominance. The second test for CPT and simplicity theory thus compares the equalizing payoffs chosen for Questions 3 and 5: CPT predicts a strictly higher equalizing payoff in Q5, while simplicity theory permits a weakly lower equalizing payoff in Q5. Note that this set of questions allows us to reject not only CPT, but also any decision theory that respects dominance.

8.2.4 Results

Table 11 summarizes the preference-elicitation results. The modal preference among our sample of binary option traders is simplicity, with 41% passing both simplicity tests and 91% passing at least one. Violations of dominance in favor of fewer-outcome lotteries are common, with 81% of respondents displaying this behavior.

Prospect theory and cumulative prospect theory preferences are markedly less present among our survey sample. About a third of the respondents pass at least one CPT test and only 3% pass both CPT tests, with 18% displaying rank dependence and 19% respecting dominance. Similarly, 50% of respondents pass the event splitting test for PT, which is again substantially smaller than the 91% that pass at least one test for simplicity.

The survey thus provides clear evidence in support of our interpretation of the market trading data. When given the opportunity to violate dominance in order to choose a lottery with fewer outcomes, the majority of our respondents do so. CPT and PT preferences – which cannot rationalize the arbitrage opportunities we observe – are significantly less common among the survey participants than a preference for simplicity.

[16]Canonical CPT refers to the commonly used formulation of Kahneman and Tversky (1992), who use the probability weighing function $w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$ and the utility function $u(x) = x^\alpha$. They identify parameter values $\gamma = 0.61$ and $\alpha = 0.88$.

Table 11: Preferences of Binary Option Traders

	Cumulative		Breakdown	
	Pass Both Tests	Pass At Least One Test	Test 1	Test 2
Simplicity	41%	91%	50%	81%
CPT	3%	34%	18%	19%
PT	-	50%	50%	-

8.3 Knowledge of Traditional Options and Access to CME

To gauge traders' knowledge of the traditional option contracts traded at CME, we asked the following question:

> Which of the following have you heard of before? (click all that apply)
>
> - Call Options
> - Put Options
> - Binary Options
> - Options on Futures (for example, options on S&P 500 index futures)
> - None of the above

Of primary interest is the fraction of survey participants who mark answer choices other than binary options (all respondents report awareness of binary options), since knowledge of call or put options and options on futures would indicate that they are familiar with the traditional option contracts we use to construct our dominating CME portfolios.

Because CME does not offer direct trading access to retail investors, respondents aware of traditional options at CME would also need an account with a licensed third-party broker to act on the arbitrage opportunities we document. To assess the extent of CME access among our survey participants, we asked whether respondents had an active trading account with any of the major retail-facing CME brokers that allow trading in options on futures:

> Do you have an account with any of the following: Charles Schwab, E TRADE, TD Ameritrade, Trade Station, Interactive Brokers, or any other licensed option brokers?
>
> - Yes
> - No

Table 12: Binary Option Traders' Access to Traditional Options

		Passed at Least One Test for:			Passed Both Tests for:	
	Overall	Simplicity	CPT	PT	Simplicity	CPT
Aware of Call or Put Options	62%	61%	58%	68%	63%	67%
Aware of Options on Futures	75%	73%	93%	78%	69%	100%
Has CME Trading Account	99%	99%	100%	100%	98%	100%

The responses to these two questions are summarized in Table 12. We report the results for the overall sample, as well as for each of the five preference classifications created by our lottery-choice questions. The results indicate that binary option traders' knowledge of traditional options is substantial: 62% are aware of call or put options, and 75% are aware of options on futures. Most respondents thus appear to have the basic institutional knowledge necessary to consider traditional option prices at CME. Trading access to CME is clearly not a limiting factor in acting on arbitrage opportunities, since 99% of the traders in the sample have an active account with a CME-licensed options broker. Institutional knowledge and CME trading access do not appear to differ substantially across the preference classifications.

8.4 Demographics and Financial Behavior

The remaining survey questions allow us to observe demographic and financial characteristics for our sample of binary option traders. A summary of the responses to these questions is given in Tables B1-B2. The traders in our sample are younger, more male, and more highly educated than the broader U.S. population. They are also notable for the amount of money they invest in the binary option market: 61% of respondents report devoting at least 40% of their risky asset holdings to binary option trading. The majority of respondents are experienced in trading binary options, as 56% have traded binary options for over two years, and 82% for at least one year.

9 Conclusion

In a first empirical study of the controversial binary option market, we use trading data to document frequent and large arbitrage opportunities created by the availability of cheaper, strictly dominant call option portfolios. Our first contribution is in documenting this anomaly; although the binary option market has generated much controversy, it has not been studied rigorously and systematic harm to retail investors has not been documented at regulated, price-transparent exchanges.

Standard explanations for the existence of arbitrage, such as random price volatility, explicit trading fees, or implicit trading costs arising from differential collateral requirements or market liquidity, cannot fully account for our results. We also reject deficient institutional knowledge as an explanation by providing survey evidence that binary option traders have the means to purchase the dominating call option portfolios we construct in our empirical analysis.

After examining standard institutional explanations, we turn to behavioral considerations. We prove that canonical revisions to expected utility theory, such as prospect theory and cumulative prospect theory, are incapable of explaining our arbitrage results. We find our results are most consistent with an alternative behavioral model in which retail traders display simplicity preferences, preferring lotteries with smaller supports and forgoing strictly dominant alternatives when they have too many possible outcomes. Direct preference-elicitation questions explicitly confirm that simplicity preferences exist and are more common than prospect theory preferences among the respondents to our survey of binary option traders.

Our findings speak directly to the controversial binary option market but also have implications for modeling and explaining household behavior in other risky financial environments with potentially complex contracts. These settings include other asset and derivative markets as well as more common household financial decisions like the choice between alternative health insurance plans. We believe the presence of simplicity preferences is relevant to academics and policymakers who study and regulate financial markets in which households participate.

References

Andreoni, J. and C. Sprenger (2011): "Uncertainty Equivalents: Testing the Limits of the Independence Axiom," Working Paper.

Barberis, N., A. Mukherjee, and B. Wang (2016): "Prospect Theory and Stock Returns: An Empirical Test," *Review of Financial Studies*, 29, 3068–3107.

Benartzi, S. and R. Thaler (2007): "Heuristics and Biases in Retirement Savings Behavior," *Journal of Economic Perspectives*, 21, 81–104.

Bernheim, B. D. and C. Sprenger (2020): "On the Empirical Validity of Cumulative Prospect Theory: Experimental Evidence of Rank-Independent Probability Weighting," *Econometrica*.

Bordalo, P., N. Gennaioli, and A. Shleifer (2013): "Salience and Asset Prices," *American Economic Review*, 103, 623–628.

Brunnermeier, M. and M. Oehmke (2009): "Complexity in Financial Markets," Working Paper.

Business Wire (2008): "The HedgeStreet Exchange Offers New Binary Options on Equity Index Futures," https://www.businesswire.com/news/home/20081120005685/en/HedgeStreet-Exchange-Offers-New-Binary-Options-Equity.

Carlin, B. I., S. Kogan, and R. Lowery (2013): "Trading Complex Assets," *Journal of Finance*, 68, 1937–1960.

Carroll, G., J. Choi, D. Laibson, B. Madrian, and A. Metrick (2009): "Optimal Defaults and Active Decisions," *Quarterly Journal of Economics*, 124, 1639–1674.

Célérier, C. and B. Vallée (2012): "Catering to Investors Through Security Design: Headline Rate and Complexity," *Quarterly Journal of Economics*, 4.

Commodity Futures Trading Commission Division of Market Oversight (2017): "Trade Practice Rule Enforcement Review: North American Derivatives Exchange," https://www.cftc.gov/sites/default/files/idc/groups/public/@iodcms/documents/file/rernadex072817.pdf.

ESMA (2019): "ESMA Ceases Renewal of Product Intervention Measure Relating to Binary Options," https://www.esma.europa.eu/press-news/esma-news/esma-ceases-renewal-product-intervention-measure-relating-binary-options.

FCA (2019): "FCA Confirms Permanent Ban on the Sale of Binary Options to Retail Consumers," https://www.fca.org.uk/news/statements/fca-confirms-permanent-ban-sale-binary-options-retail-consumers.

Fehr-Duda, H. and T. Epper (2012): "Probability and Risk: Foundations and Economic Implications of Probability- Dependent Risk Preferences," *Annual Review of Economics*, 4.

Finanstilsynet (2019): "Finanstilsynet planlægger forbud mod binære optionerfor at beskytte detailkunder," https://www.finanstilsynet.dk/Nyheder-og-Presse/Pressemeddelelser/2019/Forbud-mod-binaere-optioner-for-at-beskytte-detailkunder.

Gabaix, X. (2019): *Behavioral Inattention*, Elsevier, vol. 2, 261–343.

Graham, J. and A. Kumar (2006): "Do Divided Clinenteles Exist? Evidence on Dividend Preferences of Retail Investors," *Journal of Finance*, 61, 1305–1336.

Kahneman, D. and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47, 263–291.

——— (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

Nadex (2020): "What are Binary Options and How Do They Work?" https://www.nadex.com/products/binary-options/what-are-binary-options.

Olsen, R. (1997): "Prospect Theory as an Explanation of Risky Choice by Professional Investors: Some Evidence," *Review of Financial Economics*, 6, 225–232.

Puri, I. (2020): "Preference for Simplicity," Working Paper.

Rhode, P. and K. Strumpf (2004): "Historical Presidential Betting Markets," *Journal of Economic Perspectives*, 18, 127–171.

Sato, Y. (2014): "Opacity in Financial Markets," *Review of Financial Studies*, 27, 3502–3546.

Snowberg, E., J. Wolfers, and E. Zitzewitz (2007): "Partisan Impacts on the Economy: Evidence from Prediction Markets and Close Elections," *Quarterly Journal of Economics*, 122, 807–829.

Weinglass, S. (2019): "Israel Bans Binary Options Industry, Finally Closing Vast, 10-Year Fraud," https://www.timesofisrael.com/israel-bans-entire-binary-options-industry-finally-closing-vast-10-year-fraud/.

Wolfers, J. and E. Zitzewitz (2004): "Prediction Markets," *Journal of Economic Perspectives*, 18, 107–126.

Appendix A Partial Equilibrium Model of Retail Binary Option Market

To demonstrate how simplicity preferences on the part of retail investors can generate higher prices for dominated binary options than dominating option mixtures, consider the following partial equilibrium model.

Assume a static model: similar to a lab experiment, people first buy or sell lotteries, and then the lotteries are realized, with consumption occurring only after the realization of lotteries. Since our empirical analysis studies weekly options (a relatively short time period), this assumption seems plausible.

Let M be a large market populated primarily by professional institutional investors that is complete and where no-arbitrage holds; assume further that investors have an identical probability measure Pr over possible states of the world. Let P_B^M be the price of a binary option on underlying asset y which pays \$100 if the price of asset y exceeds K . Let P_c^M be the price of a traditional call option on underlying asset y which has strike price c. Then, by no arbitrage, $P_B^M = \lim_{\;!0} \frac{100}{}(P_K^M \quad - P_K^M)$, and $P_B^M < \frac{100}{}(P_K^M \quad - P_K^M)$ for each $\epsilon > 0$. Assume that in this market, traditional call options are traded but binary options are not (the price P_B^M just calculated is what the price would be if binary options were actually traded in this market).

Now consider a different retail market R in which only binary options are traded, and which has measure 1 of retail investors and one seller. These retail investors and seller also have probability measure Pr over possible states of the world. The price of binary option B in this market is denoted P_B^R, in contrast to the counterfactual price of the binary option in the earlier market which was denoted P_B^M. Both the retail investors and the seller have access to the earlier market[17]. The retail investors have simplicity preferences (as defined in Section 7) with $U^S(p) = \sum u(x)p(x) - C(|\text{support}(p)|)$, and the seller has expected utility preference $U(p) = \sum u(x)p(x)$, where x are monetary option payoffs. Suppose that the retail investors wish to bet on whether the price of asset y will exceed price K , that they use narrow framing, and that the only decision facing them is whether to buy traditional call options or binary options (for example, the desire to bet could come from an unmodeled portfolio choice problem or from some emotion-driven reason). By the ϵ-traditional portfolio, we mean the portfolio which buys an option on asset y with strike price $K \; - \epsilon$ and sells an option on asset y with strike price K . In this case the retail investor will choose to buy the binary option over the ϵ-traditional option portfolio if and only if, letting $q = \Pr(x > K \;)$

[17]One interpretation is that retail investors form measure zero of the larger market M, but make up all investors in the smaller retail market R.

$P^{M}_{O;} = P^{M}_{K} - P^{M}_{K}$, l the step size of the price of asset y, and $|O| = 100/l + 1$ the support size of the ϵ-traditional option portfolio,

$$\sum_{i=0}^{l} u\left(\frac{100i}{l} - P^{M}_{O;}\right) \Pr\left(x = K \;\; - \epsilon + \frac{100i}{l}\right) - C(|O|) < qu(100 - P^{M}_{B}) + (1-q)u(0 - P^{M}_{B}) - C(2). \tag{12}$$

Equivalently, the retail investor chooses the seller's binary option over the ϵ-traditional portfolio if and only if

$$\sum_{i=0}^{l} u\left(\frac{100i}{l} - P^{M}_{O;}\right) \Pr\left(x = K \;\; - \epsilon + \frac{100i}{l}\right) - qu(100 - P^{M}_{B}) - (1-q)u(0 - P^{M}_{B}) \leq C(|O|) - C(2). \tag{13}$$

Observe that the left-hand side of (13) is increasing in P^{M}_{B}. Further observe that for the seller, who can produce binary option B at cost $P^{M}_{O;l}$, if $P^{M}_{B} \geq P^{M}_{O;}$, then buying the ϵ-traditional portfolio from the larger market M, and selling the binary option in the retail market R generates a positive payoff of

$$\sum_{i=0}^{l} u\left(\frac{100i}{l} - P^{M}_{O;}\right) \Pr\left(x = K \;\; - \epsilon + \frac{100i}{l}\right) - qu(100 - P^{M}_{B}) - (1-q)u(0 - P^{M}_{B}) > 0. \tag{14}$$

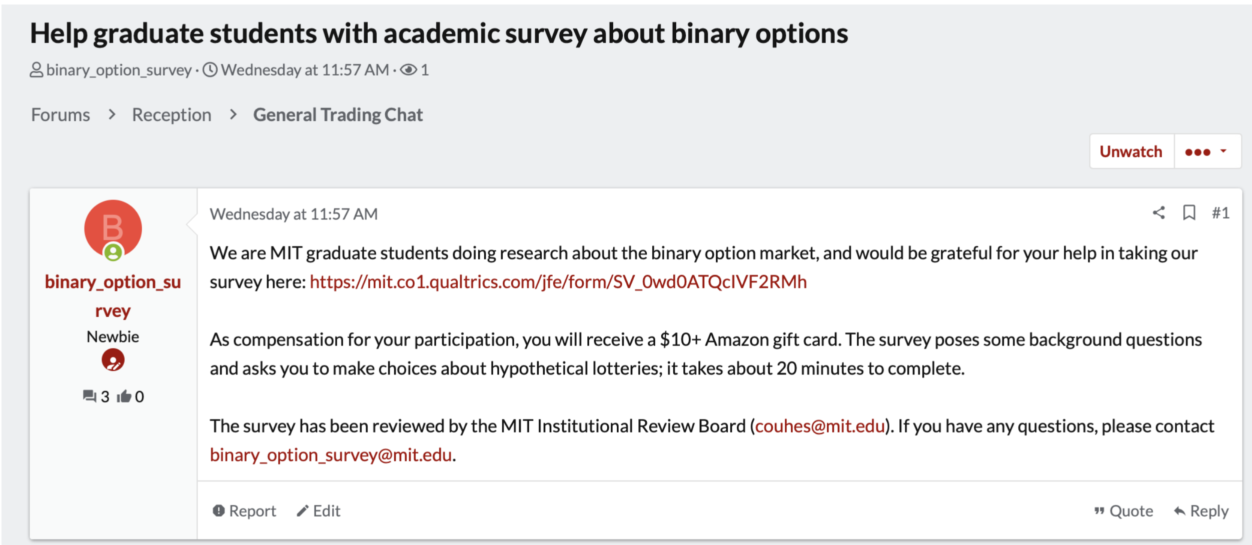
Assume a quantity constraint of Q on the number of binary options of type B which the seller can produce (for example, coming from regulatory concerns or from the software and computer memory requirements of selling too many binary options). Then the retail market R clears at Q binary options of each type sold, the price of binary option B satisfying (13) with equality, and the seller making strictly positive profit. In particular, the price of binary option B in the retail market will be strictly larger than the price of the ϵ-traditional portfolio price if either $C(|O|) - C(2)$ is large enough or if l is small enough.

Appendix B Online Survey of Binary Option Traders

B.1 Survey Posting and Questions

We advertised our survey with the following message on the "General Trading Chat" section of the trade2win.com binary option trading forum:

FIGURE B1: Online Survey Posting



The survey questions about respondent demographics and trading behavior are shown below; Section 8 shows the survey's preference-elicitation questions.

1. Have you ever traded binary options?
   - Yes
   - No
2. (If "No" to question 1): Are you interested in trading binary options?
   - Yes
   - No
3. (If "Yes" to question 1): For how long have you been trading binary options?
   - $<$ 1 year
   - 1-2 years
   - $>$ 2 years

4. What percentage of your binary option trades involve buying binary options, and what percentage involve selling binary options?
    - 75-100% of my trades are buys and 0-25% of my trades are sells
    - 50-75% of my trades are buys and 25-50% of my trades are sells
    - 25-50% of my trades are buys and 50-75% of my trades are sells
    - 0-25% of my trades are buys and 75-100% of my trades are sells
5. Which of the following have you heard of before? (click all that apply)
    - Call Options
    - Put Options
    - Binary Options
    - Options on Futures (for examnple, options on S&P 500 index futures)
    - None of the above
6. Do you have an account with any of the following: Charles Schwab, E TRADE, TD Ameritrade, Trade Station, Interactive Brokers, or any other licensed options brokers?
    - Yes
    - No
7. If you wanted to, would you find it easy to open an account with at least one of: Charles Schwab, E TRADE, TD Ameritrade, Trade Station, Interactive Brokers?
    - Yes
    - No
8. Why do you like or dislike trading binary options?
9. Approximately what percentage of your financial wealth is invested in risky assets? (Including: stocks, bonds, options, and futures; excluding: cash, checking and savings accounts, money market mutual funds)
    - 80-100%
    - 60-80%
    - 40-60%
    - 20-40%

- 0-20%

10. Of your risky asset holdings, approximately what percentage is devoted to binary option trading?

    - 80-100%
    - 60-80%
    - 40-60%
    - 20-40%
    - 0-20%

11. What is the highest level of school you have completed or the highest degree you have received?

    - Less than high school degree
    - High school graduate (high school diploma or equivalent including GED)
    - Some college but no degree
    - Associate degree in college (2-year)
    - Bachelor's degree in college (4-year)
    - Master's degree
    - Doctoral degree
    - Professional degree (JD, MD)
    - Prefer not to answer

12. What is your age?

    - 0-19
    - 20-29
    - 30-39
    - 40-49
    - 50-59
    - 60+
    - Prefer not to answer

13. Please indicate the answer that includes your entire household income in **2018** before taxes.
    - Less than $50,000
    - $50,000 to $99,999
    - $100,000 to $149,999
    - $150,000 to $199,999
    - $200,000 to $249,999
    - $250,000 to $299,999
    - $300,000 or more
14. What is your sex?
    - Male
    - Female
    - Prefer not to answer
15. Please choose one or more that you consider yourself to be:
    - White
    - Black or African-American
    - American Indian or Alaska Native
    - Asian
    - Native Hawaiian or Pacific Islander
    - Other (fill in)
    - Prefer not to answer
16. Which statement best described your employment status in **2018**?
    - Working (paid employee)
    - Working (self-employed)
    - Not working (temporary layoff from job)
    - Not working (looking for work)
    - Not working (homemaker or similar)

- Not working (retired)
- Not working (disabled)
- Not working (student)
- Not working (other) (fill in)
- Prefer not to answer

B.2 Preference-Elicitation Instructions

Figures B2-B3 show the instructions given to respondents in the preference-elicitation section of the survey.

B.3 Respondent Demographics

Tables B1-B2 summarize the demographic information provided by our survey respondents. Table B1 splits the sample with our weaker preference classifications (i.e., grouping respondents who passed at least one test for simplicity theory, cumulative prospect theory, and prospect theory), while Table B2 employs our stricter preference classifications (i.e., grouping respondents who passed both tests for simplicity theory and cumulative prospect theory).

Figure B2: Preference-Elicitation Instructions

(a)

In addition to the $10 base payment, this section gives you the opportunity to receive more money.

We will ask you about your preferences over various lotteries. **We will play out one of your choices for real** and pay you accordingly.

There are no correct answers (this is about YOUR preferences!).

It is in your financial interest to answer honestly.

(b)

EXAMPLE 1

Suppose Bob saw the following question and gives the following responses:

For the purpose of this question, LOTTERY L is:

Probability	Outcome
40 in 100 chance	$3
60 in 100 chance	$2.50

In each row, please select whether you prefer the left or right option.

For example, the first row asks whether you prefer Lottery L or $3.00 for sure.
The last row asks whether you prefer Lottery L or $1.90 for sure.

Lottery L	○	●	$3.00 for sure
Lottery L	○	●	$2.90 for sure
Lottery L	○	●	$2.80 for sure
Lottery L	●	○	$2.70 for sure
Lottery L	●	○	$2.60 for sure
Lottery L	●	○	$2.50 for sure
Lottery L	●	○	$2.40 for sure
Lottery L	●	○	$2.30 for sure
Lottery L	●	○	$2.20 for sure
Lottery L	●	○	$2.10 for sure
Lottery L	●	○	$2.00 for sure
Lottery L	●	○	$1.90 for sure

Suppose that row 11 of this question is selected for payment. In row 11, Bob indicated that he prefers Lottery L to $2. Therefore we will simulate Lottery L, and pay Bob according to the result of the simulation.

Figure B3: Preference-Elicitation Instructions Continued

EXAMPLE 2

Suppose Ann saw the following question and gives the following responses:

For the purpose of this question, "LOTTERY L" is:

Probability	**Outcome**
40 in 100 chance	$8.50
30 in 100 chance	$9
30 in 100 chance	$1.50

And "LOTTERY Q" is:

Probability	**Outcome**
40 in 100 chance	$8.50
30 in 100 chance	$9.50
30 in 100 chance	$*x*

In each row, please select whether you prefer the left or right option.

For example, the first row asks whether you prefer Lottery L, or Lottery Q with x = $1.50, and the last row asks whether you prefer Lottery L, or Lottery Q with x = $0.

Lottery L	○	●	Lottery Q with x = $1.50
Lottery L	○	●	Lottery Q with x = $1.40
Lottery L	○	●	Lottery Q with x = $1.30
Lottery L	○	●	Lottery Q with x = $1.20
Lottery L	○	●	Lottery Q with x = $1.10
Lottery L	○	●	Lottery Q with x = $1.00
Lottery L	○	●	Lottery Q with x = $0.90
Lottery L	○	●	Lottery Q with x = $0.80
Lottery L	●	○	Lottery Q with x = $0.70
Lottery L	●	○	Lottery Q with x = $0.60
Lottery L	●	○	Lottery Q with x = $0.50
Lottery L	●	○	Lottery Q with x = $0.40
Lottery L	●	○	Lottery Q with x = $0.30
Lottery L	●	○	Lottery Q with x = $0.20
Lottery L	●	○	Lottery Q with x = $0.10
Lottery L	●	○	Lottery Q with x = $0.00

Suppose that row 3 of this question is selected for payment. In row 3, Ann indicated that she prefers Lottery Q with x = $1.30 to Lottery L. Therefore we will simulate the following lottery: $8.50 with 40% probability, $9.50 with 30% probability, $1.30 with 30% probability. We will then pay Ann according to the result of the simulation.

Table B1: Demographics and Financial Behavior by Weak Preference Classifications

		By Preference Classification - Passed at Least One Test for:		
	Overall	Simplicity Theory	Cumulative Prospect Theory	Prospect Theory
Age				
20-29	26%	28%	38%	19%
30-39	65%	63%	63%	75%
40-49	7%	7%	0%	5%
50-59	2%	2%	0%	2%
Gender				
Male	73%	73%	68%	76%
Female	27%	27%	33%	24%
Race				
White	69%	68%	75%	69%
Black	19%	21%	18%	17%
Other	12%	11%	7%	14%
Education				
Some college	6%	5%	10%	8%
2-Year degree	27%	29%	28%	24%
4-Year degree	49%	48%	48%	54%
Master's degree	15%	17%	13%	12%
Doctoral degree	3%	2%	3%	2%
Employment				
Paid employee	65%	66%	65%	68%
Self-employed	28%	27%	30%	25%
Not working	7%	7%	5%	7%
Household Income				
< $50,000	3%	4%	5%	3%
$50,000-$99,999	19%	18%	23%	17%
$100,000-$149,999	25%	26%	33%	22%
$150,000-$199,999	28%	29%	18%	32%
$200,000-$249,999	15%	14%	20%	19%
$250,000-$299,999	8%	8%	3%	7%
> $300,000	1%	1%	0%	0%
Trading Experience				
< 1 year	18%	19%	13%	20%
1-2 year	26%	28%	28%	24%
> 2 years	56%	53%	60%	56%
Share of Financial Wealth in Risky Assets				
0%-20%	9%	10%	3%	10%
20%-40%	28%	29%	28%	29%
40%-60%	34%	32%	50%	31%
60%-80%	26%	26%	20%	29%
80%-100%	3%	3%	0%	2%
Share of Risky Assets in Binary Options				
0%-20%	13%	14%	3%	12%
20%-40%	26%	26%	30%	31%
40%-60%	41%	40%	40%	39%
60%-80%	17%	17%	23%	17%
80%-100%	3%	3%	5%	2%
Share of Binary Option Trades that Are Buys				
0%-25%	13%	13%	8%	12%
25%-50%	37%	36%	45%	39%
50%-75%	41%	41%	43%	41%
75%-100%	9%	9%	5%	8%
Respondents	118	107	40	59

Table B2: Demographics and Financial Behavior by Strict Preference Classifications

		By Preference Classification - Passed Both Tests for:	
	Overall	Simplicity Theory	Cumulative Prospect Theory
Age			
20-29	26%	29%	33%
30-39	65%	58%	67%
40-49	7%	10%	0%
50-59	2%	2%	0%
Gender			
Male	73%	71%	33%
Female	27%	29%	67%
Race			
White	69%	71%	100%
Black	19%	21%	0%
Other	12%	8%	0%
Education			
Some college	6%	2%	33%
2-Year degree	27%	29%	0%
4-Year degree	49%	44%	67%
Master's degree	15%	21%	0%
Doctoral degree	3%	4%	0%
Employment			
Paid employee	65%	63%	33%
Self-employed	28%	29%	33%
Not working	7%	8%	33%
Household Income			
< $50,000	3%	4%	0%
$50,000-$99,999	19%	19%	67%
$100,000-$149,999	25%	25%	0%
$150,000-$199,999	28%	25%	0%
$200,000-$249,999	15%	13%	33%
$250,000-$299,999	8%	13%	0%
> $300,000	1%	2%	0%
Trading Experience			
< 1 year	18%	19%	0%
1-2 year	26%	27%	33%
> 2 years	56%	54%	67%
Share of Financial Wealth in Risky Assets			
0%-20%	9%	8%	0%
20%-40%	28%	23%	33%
40%-60%	34%	38%	33%
60%-80%	26%	27%	33%
80%-100%	3%	4%	0%
Share of Risky Assets in Binary Options			
0%-20%	13%	15%	0%
20%-40%	26%	23%	0%
40%-60%	41%	40%	67%
60%-80%	17%	19%	33%
80%-100%	3%	4%	0%
Share of Binary Option Trades that Are Buys			
0%-25%	13%	15%	33%
25%-50%	37%	33%	33%
50%-75%	41%	42%	33%
75%-100%	9%	10%	0%
Respondents	118	48	3