Three Essays in Applied Economics

On Exploiting Arbitrage and Detecting In-Auction Fraud in Online Markets

Dissertation
zur Erlangung der Würde eines
Doktors der Staatswissenschaften

vorgelegt der
Wirschaftswissenschaftlichen Fakultät der
Universität Basel

von Dominic Herzog
von Basel, Basel-Stadt

Buchbinderei Beat Gschwind
Basel, 2015

Genehmigt von der Wirtschaftswissenschaftlichen Fakultät der Universität Basel auf Antrag von Prof. Dr. Yvan Lengwiler und Prof. Dr. Dietmar Maringer.

Basel, den 27. Januar 2015 Der Dekan: Prof. Dr. Yvan Lengwiler

Acknowledgments

I look back on the past four years as having given me a once-in-a-lifetime opportunity to commit myself completely to researching the questions that my studies had aroused. During this time, numerous people accompanied and supported me in different ways, and I truly want to thank them for their encouragement. To my thesis advisor Yvan Lengwiler, I express my deepest gratitude for his unambiguous support and his strong backing of my unusual research ideas as well as the freedom granted to pursue these objectives. His exceptional ability to grasp all kind of ideas and discuss them in the light of his profound economic knowledge resulted in many lively and fruitful discussions. In the first chapter, he pointed out that political betting targets a wider audience and might, therefore, be preferred to sports betting. In addition, he personally appealed for the monetary support required for the field experiment. I owe Yvan Lengwiler a debt of gratitude for providing me the opportunity to write this thesis under his supervision.

Sincere thanks to Dietmar Maringer, who agreed to support my thesis as second advisor without hesitation, for his insightful comments. I truly thank Hermione Miller-Moser for her linguistic corrections and annotations to this thesis.

I thank three anonymous referees for their helpful comments on the first chapter as well as an anonymous referee for his inputs on the second chapter. Furthermore, I cordially thank all my colleagues at the department who accompanied me during my doctoral studies. I really appreciate all the remarks, thoughts and constructive criticism that they offered in interesting discussions.

I thank my family for their boundless support. I know that you will always be there for me. Finally, I owe gratitude to my loving wife Fabienne. She is my anchor, my everlasting drive to develop further and provides me with unconditional support in achieving my aims.

Introduction

This thesis contains three essays in applied economics which all deal with online markets. The basic idea behind my investigations might seem simple: Identifying valuable information and transforming the obtained advance knowledge into monetary profit, or, at least, into fresh market insights.

Online markets provide favorable conditions for implementing automated tasks which would otherwise be highly time-consuming. By automatically recording and analyzing online market events, I am able to acquire valuable information regarding a large number of current interactions with a reasonable amount of effort. The acquired datasets provide a unique opportunity to investigate the market participants' behavior. However, the three chapters of my thesis are not solely concerned with analyzing behavioral data. They are also about recognizing profitable market opportunities that arise, for example, from loopholes in the market's design or from information which should not be public, and acquiring the unique datasets which are the empirical base of the thesis papers. Therefore, Identifying, collecting and evaluating valuable information for this thesis in real-time in order to exploit current bargain opportunities is a challenging and rewarding task.

The first chapter addresses the online betting market with regard to the 2012 U.S. Presidential Election. In particular, I recorded minute-by-minute odds from dozens of bookmakers and betting exchanges over a period of 14 months prior to the election. By immediately analyzing this stream of information, I was able to identify price misalignments across the market in real-time. In order to demonstrate that the observed violation of the law of one price actually provides exploitable arbitrage opportunities, I invested real money according to an arbitrage-based strategy. This strategy excludes all potential risks, except for the risk that a platform becomes insolvent. The field experiment highlights the market's accessibility and the costs incurred by implementing the strategy and placing the investments. My methodology therefore enables me to provide evidence that exploitable inter-market arbitrage opportunities exist in this market, affords an insight into the market's dynamics, and also allows me to enjoy a free lunch.

Chapter 2 investigates the pay-per-bid auctions provided by labuyla.ch. The disclosure of the auctioned item's hidden price is the main incentive for placing a bid in price reveal auctions. However, I observed that a loophole existed in labuyla.ch's platform which allowed an attentive observer to calculate the hidden price for free. In addition, their auction design allowed me to observe and record the bidding behavior of all the participants on their platform. The analysis of the dataset reveals that this loophole was unknown to the active bidders. Moreover, my investigations on the observed behavior lead to the sole conclusion that the auctioneer himself was cheating by participating in the auctions as seller, bidder and buyer. I confronted the owner of labuyla.ch with my research results and my findings were confirmed by the fact that they rapidly shut down all of their auctions after our discussion.

In Chapter 3, I deal with another in-auction fraud, namely: shill bidding. Shill bidding describes the fraudulent behavior of a seller who bids in his own auctions. The anonymity in online auctions and the fact that a single bidder is allowed to open up several accounts on a platform favors such misbehavior. Since auction houses profit from higher sales prices, they have a reduced incentive to prohibit shilling. Several approaches in the literature tackle this problem by identifying shill bidders on the basis of their publicly observable behavior. While inspecting ricardo.ch's website for public behavioral data, I detected an information leakage — now closed — that allowed me to observe all accounts' personal details as well as the entered bidders' valuations of the auctioned item. During a four-month period I recorded the bidding history of nearly two million auctions. By comparing the seller's and the bidders' personal details (name, address, and phone number) in these auctions, I was able to accurately identify shill bidders in my dataset and analyze their behavior. In addition, I test the accuracy of two identification algorithms which are based on public information.

Contents

List of Figures

List of Tables

Chapter I

Exploiting Arbitrage Opportunities in Online Betting: Empirical Evidence from the 2012 U.S. Presidential Elections

Abstract. This paper addresses the observation of arbitrage opportunities across online betting markets and the empirical evidence of their exploitation. During the last 14 months prior to the 2012 U.S. Presidential Election, I extracted minute-by-minute odds from numerous bookmakers and betting exchanges. I show that persistent price disparities existed on the election outcome's odds which allowed for implementing a risk-free investment strategy. By placing real investments, I was able to identify the actual costs and risks of this strategy and empirically prove that exploitable arbitrage opportunities exist in inter-market online betting.

Keywords: arbitrage; online betting; bookmaker; prediction market; betting exchanges; election

Jelclass: D49; G14

1 Introduction

In this paper, I provide empirical evidence that exploitable inter-market arbitrage opportunities exist in the online betting market. The empirical evidence results from real investments which I made in order to exploit odds disparities on the 2012 U.S. Presidential Election outcome across bookmakers and betting exchanges. The applied arbitrage-based strategy does not, for example, face the risk that a platform restricts or closes the investor's account. My suggested strategy ensures that the placed bets result in the same profit no matter what the event's outcome is or whether a bookmaker refuses future bets. By personally placing bets, I am able to quantify the costs, assess the risks and identify barriers to the market which might limit the exploitation of such abnormal returns and might explain their existence. Therefore, my findings provide consistent evidence on the successful exploitation of theoretically observed inefficiencies, which, for example, Oikonomidis and Johnson [2011] call for in their paper.

Several studies, which I will discuss in more detail in the next section, report the existence of arbitrage opportunities as a result of analyzing odds disparities on online betting markets. However, so far, very few papers were able to ascertain the monetary value of the gains that can be drawn from exploiting the identified arbitrage opportunities. One reason is that their findings are based on past events, and, therefore, a comprehensive analysis of whether or not their strategies can be implemented is limited. Such studies find evidence of what I call surebet opportunities. Surebets result from odds disparities across different gambling markets and imply that an investor can gain a positive profit by simultaneously betting on each possible outcome of an event. However, a surebet is not in itself evidence of arbitrage. To ascertain that arbitrage is possible, all the costs, risks, barriers and limits involved in an investment must also be considered.

My investigations are based on minute-by-minute odds which I recorded from dozens of bookmakers and betting exchanges over a period of 14 months prior to the 2012 U.S. Presidential Election. By observing the currently available odds and instantly checking these values for arbitrage opportunities, I show that the information on currently existing weak-form market inefficiencies is available. Moreover, the minute-by-minute data allow me to assess how persistent such exploitable inter-market arbitrage opportunities are.

By analyzing the offered odds on numerous platforms with regard to exploitable arbitrage, I aim to present the online betting market in its entirety. If individuals place investments at more than one platform, then observing only one leg of their activities might lead to a false assessment of their behavior. The inter-market focus of this paper aims to discover strategies which involve activities on more than one platform and, therefore, provide fresh insights on former observed investment patterns. Moreover, exploitable cross-market arbitrage might prevent market prices from swiftly reacting to new information if these platforms adjust prices at different rates. If, however, arbitrageurs are present, then price manipulation is limited as well and becomes more costly to maintain. Arbitrageurs might also link betters who do not have access to all platforms and thereby exchange information and beliefs.

The remainder of the paper is structured as follows. Section 2 discusses the related literature and highlights the contribution of this paper. Section 3 characterizes the online betting market, describes the extracted data, and explains my devised arbitrage-based strategy. Section 4 analyzes the 2012 U.S. Presidential Primary Elections with regard to surebets. In Section 5, I describe and discuss the experiment's findings. Section 6 compares the months leading up to the election with the observations made during the primaries. Finally, Section 7 concludes.

2 Literature

This paper was inspired by Levitt's [2004] article, 'Why are gambling markets organized so differently from financial markets'. Besides the markets' similarities (e.g., investors with heterogeneous beliefs or the markets' zero-sum game structure), Levitt highlights the markets' differences. With regard to gambling markets, Levitt discusses bookmakers' strategies. He concludes that bookmakers only hire employees who are excellent at predicting a game's outcome in order to generate winnings. In addition, bookmakers use their skills to deviate from the odds that clear the market in order to exploit bidders biases and consequently achieve higher profits. However, asset quotes on financial markets are not determined by a few individuals. Financial markets serve to match the demand and supply of many investors, which results in a continuously changing price.

Over the past years, prediction markets for sporting and political events have developed alongside traditional bookmakers services. The final cash value of their traded assets depends on a specific event's outcome. Prediction markets are organized like financial markets, but operate in the market traditionally covered by bookmakers. These days, gambling markets (bookmakers) and financial markets (prediction markets) offer binary options on the same events and, therefore, the law of one price should hold.

Levitt's finding that bookmakers intentionally deviate from the market clearing price might also open up arbitrage betting opportunities in the market. In this context, the most widely discussed form of bias is the favorite-longshot bias (Griffith, 1949; Snowberg and Wolfers, 2010). This behavioral bias describes a strategy which systematically undervalues (odds are too high) favorites and systematically overvalues underdogs (odds are too low). Other reported betting market biases are, for example, anti-incumbency (Rothschild, 2009), the home bias (Chincarini and Contreras, 2010), the disposition effect (Borghesi, 2011), the reverse favorite-longshot bias (Woodland and Woodland, 2003), and market overreaction (Tetlock, 2004).

Vlastakis *et al.* [2009] as well as Franck *et al.* [2013] provide a brief overview of the recent literature on betting market efficiency. As most of these studies (e.g., Hausch and Ziemba, 1990; Dixon and Pope, 2004; Edelman and O'Brian, 2004; Paton and Vaughan Williams, 2005) use data on past events, their findings refer to surebet opportunities. However, risks (e.g., counterparty risk), costs (e.g., transaction costs), and market access, as well as the availability of the necessary information when the opportunity occurred must also be considered in order to conclude that arbitraging was actually possible. Chincarini and Contreras [2010] refer to the currency risk involved and note that if all accounts were run in U.S. dollars, then this risk would be eliminated. Franck *et al.* [2013] and Tetlock [2004] discuss the bet's counter-party risk. In a recently published paper, Rothschild and Pennock [2014] investigate price misalignment in prediction markets and find a persistent price spread between Intrade and Betfair. They address in detail the costs and risks of a potential arbitrage strategy and discuss how the observed price misalignment can persistently exist. In addition, they invested a few thousand dollars in arbitrage bets in order to assess the size of the shadow order book in these exchanges.

Like Franck *et al.* [2013], I focus on inter-market arbitrage opportunities in online betting markets, and consider bookmakers as well as prediction markets. Franck *et al.* [2013] refer to data on past European soccer games, which cover a total of 11,933 top-five league matches over seven seasons (from 2004/05 to 2010/11). They find evidence that: "For 19.2% of the matches, the return on the short position hedged bet is positive and yields an arbitrage opportunity with an average positive return of 1.4%." [Franck *et al.*, 2013, p. 308–309] Their study investigates liquidity constraints and fees, but as their analyses lack comprehensive data on these issues, they are unable to accurately estimate the degree of market liquidity.

Croxson and Reade [2014] find that prices on Betfair swiftly and fully process new information. Their analyses are based on second-by-second snapshots of Betfair's order book on soccer-related markets. To investigate the market's information processing, they analyze price developments in matches where goals were scored on the cusp of the newsless half-time break.

Gil and Levitt [2007] analyze Intrade's full order book for the soccer 2002 World Cup second-by-second and find mixed support for the efficient market hypothesis. Their dataset also contains the investors identification number which allows them to identify the market makers who provide liquidity in these markets.

Rothschild and Sethi [2014] provide a comprehensive analysis of individuals' investment behavior on Intrade, covering the entire two-year trading period prior to the 2012 U.S. Presidential Election. Based on transaction-level data provided by Intrade, they are able to ex-post identify and analyze individuals' trading activities. They assign individuals to one of the six categories of trading strategy that they identify based on six behavior characteristics. Rothschild and Sethi [2014, p. 22] find "that the trading process is driven, in large measure, by individuals with different *interpretations* of *public* information." In addition, Rothschild and Sethi [2014, p. 4] identify one investor who manipulated the market prices during the election day "by placing large bids for Romney and large orders for Obama that effectively created a firewall, preventing prices from moving in response to incoming information." While Rothschild and Sethi [2014] analyze the complete transaction level data of all investors on Intrade, I focus on price comparisons across different platforms. This comparison has the potential to provide evidence that other investors, besides me, who are classified as 'Extreme Bias' by Rothschild and Sethi [2014] were actually exploiting cross-market arbitrage.

Widely discussed price biases in the betting market might provide an opportunity to gain abnormal returns in the long run (e.g., Lane and Ziemba, 2004; Paton and Vaughan Williams, 2005; Borghesi, 2007) and might influence the reliable prediction of an event's outcome based on betting odds (e.g., Spann and Skiera, 2009; Snowberg *et al.*, 2012; Rothschild, 2009). However, neither of these two applications are within the scope of this paper. Strategies which try to exploit biases through positive expected returns in the long run face the risk that the bookmaker will exercise his right to restrict or close the investor's account. Several studies (e.g., Franck *et al.*, 2013) refer to arbitrageurs who were affected by such sanctions, and, therefore, highlight this uncertainty. For this reason, I focus on surebets which yield the same profit for all possible outcomes of an event and which are unaffected by sanctions that might restrict an investor's future bidding options.

In this paper, price biases only become relevant if they differ across platforms and, thereby, open up surebet opportunities. Such differences across platforms are, for example, described by Smith *et al.* [2006] who find evidence that betting exchanges exhibit a significantly lower market bias compared to bookmakers.

3 Materials and Methods

This section aims to provide the needed background in order to investigate this paper's hypothesis:

Hypothesis. *Inter-market online betting offers exploitable arbitrage opportunities.*

Firstly, I discuss overlaps and differences between bookmakers and betting exchanges, briefly explain their functionality, and define how I handle the diversity of provided odds formats. In Subsection 3.4, I outline my suggested arbitrage-based strategy to exploit price misalignments. Subsection 3.5 describes the covered events and defines the two different periods under investigation. Finally, Subsection 3.6 provides information on the acquired and analyzed dataset.

3.1 Odds Formats

Identical odds can be expressed in different ways. Bookmakers and betting exchanges typically use English or European odds. Suppose that a bet on an event's outcome has European odds of 5. This means that if the outcome occurs, then the bettor receives five dollars back for every one dollar at stake. Therefore, European odds express how many units will be paid out per unit staked. The equivalent English odds would be 4/1 and implies that the bettor wins four dollars for every dollar at stake. If the outcome occurs, then the bettor receives his four-dollar winnings and his one-dollar stake back. Therefore, English odds express how many units you win per unit staked, which is equivalent to the return on investment (ROI). However, some betting exchanges (e.g., Intrade) determine an asset's final values, which depend on whether or not the underlying event's outcome occurs. These binary options are then traded among investors, while the asset's price in relation to the winner's payoff can be interpreted as the probability that the outcome will occur.

To increase comparability within this paper, I represent all odds and asset prices as the implied probability ρ that a specific outcome will occur. The relationship across different formats is:

$$\rho = \frac{1}{\text{English Odds} + 1} = \frac{1}{\text{European Odds}} = \frac{\text{price}}{\text{winner's payoff}}.$$

The subscripted character specifies whether the probability of occurrence is implied by an available back (ρ_b) or lay (ρ_l) offer. To back a bet is equivalent to taking a long position by buying contracts in the financial market, or to bet on an outcome's occurrence, respectively. On the contrary, laying a bet is equivalent to taking a short position by selling contracts in the financial market, or to bet against an outcome's occurrence, respectively.

3.2 Bookmakers

Bookmakers determine by themselves which events they are willing to offer a limited stack of odds on. Moreover, they also set each investor's betting limits. For the investor, the odds are take-it-or-leave-it offers. However, the displayed odds are without obligation for the bookmaker, because each bookmaker analyzed in this paper has a clause in its terms and conditions which is equivalent to WilliamHill's paragraph 9.4: "We reserve the right to refuse the whole or part of any Transaction requested by You at any time in our sole discretion, or where You have breached the Terms of Use."

With a bookmaker, the investor can only bet on an outcome occurring, because the bookmaker always bets against the outcome's occurrence. The bookmaker loses on some outcomes (exactly one if odds are offered on all possible outcomes and these outcomes are disjoint), and wins on all others. Bookmakers can sustain profits by overrounding their balanced book. In the literature (e.g., Franck *et al.*, 2013), this overround, which measures the bookmaker's share as a percentage of the payout, is interpreted as the bookmaker's commission. However, I define the bookmaker's commission according to the commissions taken by Betfair and WBX: as a percentage of the investments made. This definition of the bookmaker's commission is called vigorish and equals $\frac{\text{overround}}{1+\text{overround}}$.

Example 1 The fair European Odds on each outcome (heads or tails) of a coin toss is 2, which implies a winning probability of $\frac{1}{2} = 50\%$ per outcome. Suppose now that a bookmaker offers odds of only 1.6 on each outcome. The book is balanced if the total amount invested by individuals in each outcome 'heads' and 'tails' is equal. Odds of 1.6 imply a winning probability of $\frac{1}{1.6} = 62.5\%$ for each outcome and a total winning probability of $2 \cdot \frac{1}{1.6} = 125\%$, or an 25% overround, respectively. In this case, the bookmaker has a sure profit on his balanced book due to his vigorish of 20% on the total investment.

Because bets are zero sum games, the vigorish equals the negative weighted ROI value across all investors.

3.3 Prediction Markets

In contrast to bookmakers, prediction markets provide a service that matches the demand and supply for wagers by means of a double auction. Any investor can offer back as well as lay positions. Moreover, he can freely choose the odds and the amount of money of his order. All orders are listed in the book, and whenever a seller offers a price which a buyer accepts, a trade occurs. To guarantee that an investor is able to bear all possible losses, he must have sufficient free funds on his user account. This amount is frozen until the bet is settled or the investor withdraws an unmatched order. As previously mentioned, some prediction markets restrict investors to the trading of standardized assets which pay a fixed amount to the winner.

> **Example 2** On Intrade, an asset for a specific outcome of an event will pay the winner USD 10. Suppose that an investor is willing to buy one unit of this binary option at a price of USD 6. Another investor offers a lay position at the same price and is thus willing to pay USD 4 in order to receive a payoff of USD 10 if the outcome does not occur. The provider matches these investors and freezes their possible losses, USD 4 on the buyer's account and USD 6 on the seller's account. This trade implies a probability of $\frac{\text{price}}{\text{winner's payoff}} = \frac{\text{USD } 6}{\text{USD } 10} =$ 60% that the underlying outcome occurs.

Prediction markets make their profits either by charging investors a fixed fee for using their service or by charging a commission on net winnings.

3.4 Surebet and Arbitrage

Bodie *et al.* [2011, p. 324] define arbitrage as follows: "An arbitrage opportunity arises when an investor can earn riskless profits without making net investments." In order to accommodate the expression of arbitrage, I define a concept called a 'surebet'. A surebet implies possible price inefficiency in the market. Such inefficiencies arise whenever an investor can simultaneously place bets which, as a whole, yield the same positive profit for each possible event's outcome. However, the fact that candidates might switch parties during the election campaign, as Gary Johnson did on December 28, 2011, highlights the importance of the same underlying event.

Surebets can be constructed in various ways. Placing a bet with a bookmaker and selling the same bet at a prediction market is such an example. Franck *et al.* [2013] call this strategy, which requires only two bets, short-position inter-market arbitrage. They found that this surebet construction generates the highest returns. A surebet according to this strategy occurs whenever $\rho_b < \rho_l$ holds for the same event's outcome. An equal profit for either outcome can be reached by investing $i_{lay} = (1 - \rho_l) \cdot 10\$$ and $i_{back} = \rho_b \cdot 10\$$. This strategy gains a sure profit of $\pi = (\rho_l - \rho_b) \cdot 10\$$, implying a ROI of $\frac{\rho_l - \rho_b}{1-(\rho_l - \rho_b)}$.

> **Example 3** Suppose, that a bookmaker offers the (European) odds of 8 on a specific outcome. At the same time, a prediction market (offering assets which pay USD 10 to the winner) has an unmatched back offer at USD 2.86 for the same event's outcome. An investor can match this offer by laying at USD 2.86. This investment costs him USD 7.14 and results in a payoff of USD 10 if the outcome does not occur (zero elsewhere). Simultaneously, the investor can back USD 1.25 against the bookmaker, which costs him USD 1.25 and results in a payoff of USD 10 if the outcome does occur (zero elsewhere). The total investment made is USD 8.35, and the sure payoff is USD 10. The ROI of this surebet is 19%.

The prediction markets' commission on net winnings within a market can be considered by adjusting the implied winning probabilities. Such an adjustment, where the lay probability equals $1 - \frac{1-\rho_l}{1-\text{commission}\cdot\rho_l}$ and the back probability equals $\frac{\rho_b}{1-\text{commission}}$, can be interpreted as a higher investment that outweighs the exchanges' commission.

I discuss the potential of some further surebet strategies in Section 4 and Section 6. If not stated otherwise, the term surebet refers to the strategy described above. A surebet is a necessary but not a sufficient condition for pure arbitrage in the online betting market which always results in the identical profit. Potential counterparty- and currency risks, transaction costs, or the fact that the invested money is frozen until the event occurs are explanations for observing imaginary arbitrage. From my point of view, the term arbitrage is often hastily used in the literature, even if the attenuated label arbitrage opportunity is correct. In order to state that arbitrage is actually possible, all these costs and risks must be considered. However, as long as this is not the case, I use the term surebet.

I do not state that my suggested investment strategy is optimal. According to the investor's preferences, he should probably take at least some risk in order to increase his expected profit. However, exploitable surebets show that riskless profits are achievable and, therefore, the expected profit of risky behavior should be compared to the gains of these surebets.

3.5 Covered Events

In order to handle dozens of providers as well as a high frequency of observations, I needed to limit the analyses to a few selected events. These events had to meet the following criteria: (1) Bets on these events had to be offered by numerous bookmakers and prediction markets in order to allow an inter-market comparison; (2) The events had to be of sufficient public interest to secure adequate liquidity on the betting exchanges. This liquidity is necessary to generate sufficiently high investments that are capable of covering the fixed costs of the suggested strategy; and, (3) The corresponding bets had to be available sufficiently far in advance in order to gain insight into the persistency of potential arbitrage opportunities.

The 2012 U.S. Presidential Election (including the Primaries) fulfilled these criteria. In the period between August 27, 2011 and November 07, 2012, this paper covers four different events of the 2012 U.S. Presidential Election, namely: (1) 2012 Presidential Election Winner (Individual); (2) 2012 Presidential Election Winner (Party); (3) 2012 Democratic Presidential Nominee; and (4) 2012 Republican Presidential Nominee. The outcome of a specific event is, for example, labeled as '1:Obama', which means that Barack Obama wins the 2012 U.S. Presidential Election.

I divided the period into two phases, entitled: 'primaries' and 'runoff'. The primaries lasted from August 27, 2011 until May 11, 2012. During this phase I observed and analyzed the online betting market in real-time with regard to surebets. The results are described in Section 4. My behavior during the runoff, from May 12, 2012 until November 07, 2012, was based on the experience that I gained from the primaries. I also consider the Event 3 and the Event 4, but only for analyses during the primaries. Actually, the official primaries lasted until the announcement of the official candidate nomination at the National Convention of the parties. Mitt Romney was officially nominated as the Republican Presidential Candidate on August 29, 2012. Barack Obama was officially nominated as the Democratic Presidential Candidate on September 05, 2012. Quotes on Event 3 and Event 4 were offered until these dates. However, Romney's and Obama's nomination were already certain on May 12, 2012. Information about the covered platforms and their offered events is given in Table 1.

3.6 Origin of the Data

To discover current surebets in ongoing bets, I wrote a program to extract the current quotes from four prediction markets, and dozens of bookmakers. All outcomes of the four events were covered. The automated extraction occurred every 60 seconds, such that any change in the odds can be dated very precisely. The data were analyzed regarding surebets and stored in a database. Finally, a summary of the current surebet opportunities was periodically sent to my mobile phone. The extracted observations contain information about the: platform, event, outcome, timestamp, quote format and the odds for bookmakers, or the order book (up to 15 positions) for prediction markets. The timestamp of each quote was recorded exactly to the second. However, I round the observation time up to the next minute in order to compare different providers' odds.

The high frequency of data observations and the numerous platforms covered in this paper permit a, so far, unique insight into the dynamics of the betting market. In addition, the analysis of currently available back and lay positions at numerous bookmakers and prediction markets allows me to identify current surebets opportunities.

4 Primaries: Market Analyses

By observing and analyzing current quotes during the primaries, I examine two necessary conditions regarding my hypothesis. Firstly — and even though the existence of surebets

Table 1: Overview of Bookmakers and Prediction Markets

	Provider		Primaries				Runoff	
	Name	URL	1	2	3	4	1	2
Bookmaker	10Bet[4]	10bet.com					x	x
	888sport	888sport.com					x	x
	Bet365	bet365.com	x	x		x	x	x
	Betfred[3]	betfred.com					x	
	Betsson[1]	betsson.com					x	
	BlueSquare[2]	bluesq.com		x		x	x	x
	Bodog	bodog.co.uk	x				x	
	Boylesports[2]	boylesports.com					x	
	Bwin[1]	bwin.com					x	x
	Centrebet[2]	centrebet.com					x	x
	Coral[2]	coral.co.uk					x	x
	Intertops	intertops.eu	x	x		x	x	x
	Ladbrokes	ladbrokes.com	x				x	
	PaddyPower[3]	paddypower.com	x	x	x	x	x	x
	PinnacleSports	pinnaclesports.com		x				x
	SkyBet[2]	skybet.com	x	x			x	x
	Sportingbet	sportingbet.com	x	x	x	x	x	x
	StanJames[5]	stanjames.com	x	x	x	x	x	
	Totesport[3]	totesport.com					x	
	Unibet[1]	unibet.com					x	x
	VictorChandler[5]	betvictor.com	x	x	x	x	x	x
	WilliamHill	williamhill.com	x	x		x	x	x
Exchanges	Betfair	betfair.com	x	x	x	x	x	x
	IEM[6]	tippie.uiowa.edu/iem		x		x		x
	Intrade	intrade.com	x	x	x	x	x	x
	WBX	wbx.com	x	x	x	x	x	x

Notes: This table provides an overview of the bookmakers and prediction markets considered in this paper. Besides the providers' website URL, from which I retrieved the current quote, the covered events during the primaries and runoff are marked. No 'x' during the primaries implies that I observed the quotes of this bookmaker only during the runoff. The meaning of the superscripts is as follows:

[1] Deposit of money is not free of charge.

[2] Betting services are not available for Swiss residents.

[3] The user account cannot be administered in U.S. dollars.

[4] A minimum turnover of the deposited money is required.

[5] The information on the investor's individual limit is not provided by the host prior to a deposit.

[6] Settlement is based on the popular vote plurality and not on who is elected.

has already been confirmed in several studies — I show that surebet opportunities exist in inter-market online betting.

Requirement 1. *Inter-market online betting offers surebet opportunities.*

Secondly, I require that the information needed to identify a surebet is available for ongoing bets and that an investor has access to the corresponding betting platforms.

Requirement 2. *Exploitable surebets exist in inter-market online betting.*

This section is structured as follows. Subsection 4.1 and Subsection 4.2 briefly describe the covered platforms. Subsection 4.3 highlights the providers' importance with respect to Requirement 1. Subsection 4.4 and Subsection 4.5 analyze the market in order to identify surebets and capture their characteristics. Subsection 4.6 incorporates my observations on the market developments within a couple of preliminary results and motivates my suggested arbitrage-based strategy in the field experiment.

4.1 Bookmakers at the Primaries

My market observation during the primaries covered a total of 12 bookmakers, namely: Bet365, BlueSquare, Bodog, Intertops, Ladbrokes, PaddyPower, PinnacleSports, SkyBet, Sportingbet, StanJames, VictorChandler and WilliamHill.

Instead of charging an explicit commission, a bookmaker is sure to obtain the vigorish by balancing his overrounded book. However, a bookmaker faces a trade-off between obtaining higher profits on his balanced book (increasing the vigorish) and attracting more investors (lowering the vigorish). The second option bears higher risks for the bookmaker, because if his current odds do not balance the book, he might be forced to adjust them. However, such an adjustment might result in a monetary loss in at least some event's outcomes. Therefore, the bookmakers must have the skills to reflect and respond to changes in investors' beliefs. This fact is even more important if his attractive odds allow for inter-market surebets.

4.2 Prediction Markets at the Primaries

I observed and recorded the order book of four different prediction markets, namely: IEM (Iowa Electronic Markets), WBX (World Bet Exchange), Intrade, and Betfair. The data contain the 1 (IEM), 3 (Betfair), 5 (WBX) and 15 (Intrade) best offers on each side of a double-auction.

Market liquidity is an essential condition with regard to the possibility of investing significant amounts of money in order to cover the strategy's fixed costs. The IEM is a small-scale futures market which is run by the University of Iowa, primarily for the purpose of research. For this reason, IEM limits an individual's total investments to USD 500 and is the only U.S. prediction market which is tolerated by the Commodity Futures Trading Commission (Arrow *et al.*, 2008). In addition, the settlement of IEM's winner-takes-all market is based on the popular vote plurality winner and not on the elected President's party. The 2000 U.S. Presidential Election has shown that this difference in the settlement rule can be important, because at that election the Democrats won the popular vote plurality, but a Republican was the elected President (George W. Bush). Thus, this difference prevents the possibility of making a direct odds comparison between IEM and other providers, and I therefore exclude IEM from all further analyses.

On WBX, the liquidity was so low that in over 50% of the observations no open lay offer existed. Moreover, if the market had an open position, then the average market size on WBX was about one-fifth of its size on Intrade or Betfair. Intrade and Betfair are similar with regard to the bets they offered and their market size. Even though Gorge W. Bush signed the Unlawful Internet Gambling Enforcement Act in 2006 (Arrow *et al.*, 2008), Intrade's services were also available to U.S. citizens until the end of 2012. However, Betfair's services are classified as games of chance and are thus not available to U.S. citizens.

Intrade charges a fixed monthly fee of USD 4.99 for their services. During the 2012 U.S. Presidential Election, Betfair and WBX charged a 5% commission on net winnings within a market. A discount system might reduce the commissions that Betfair and WBX require in stages down to 2.5% and 2%, respectively. However, to reach such a low level, the investor must have paid commissions amounting to, at least, USD 12,600 on Betfair, or USD 9,200 on WBX, respectively. Whenever I use prices offered by Betfair or WBX for calculations, I take the maximum possible commission of 5% into account.

4.3 Efficiency and Vigorish

Bookmakers as well as prediction markets should be, at least if they are considered individually, efficient. If not stated otherwise, I use the term 'efficient', signifying that the currently offered prices do not allow for arbitrage by simultaneous betting on several event's outcomes.

A bookmaker who sets odds which imply a negative overround, or an inefficient prediction market, would provide an easy opportunity to gain risk-free profits. The first column of Table 2 reports the bookmakers' vigorish and the prediction markets' efficiency. The calculated vigorish values and efficiencies refer to Event 2 and consider only the outcomes 2:Democrats and 2:Republicans. In order to increase comparability across platforms, I do not include the odds on independent parties running for election, because the odds on 2:Independent are rarely offered. The vigorish values presented are larger than zero for every single observation, and would be even larger if the outcome 2:Independent had been included. No comparable vigorish can be calculated for Bodog and Ladbrokes, which do not offer bets on the winning party.

The prediction markets' efficiency, also in the first column of Table 2, is calculated according to the bookmakers' vigorish by using the back positions' odds. However, these efficiency values result from market behavior, whereas bookmakers set the vigorish.

A low vigorish implies that the bookmaker is offering, at least for one outcome, comparatively higher, and thus more attractive, odds. The bookmakers' average vigorish during the primaries lies between 6.9% and 8.6%, and is, thus, slightly higher than the previously reported value of 5% (Gil and Levitt, 2007; Rhode and Strumpf, 2004) for a typical bookmaker. Including odds for an independent party where possible increases the observed vigorish by 1.3% on average. Since PinnacleSports has a much smaller vigorish of 2.2%, I pay special attention to this bookmaker in the following analyses.

The values for the efficiency on Betfair and WBX are statistically significantly larger than zero. In contrast, Intrade's value is statistically significantly negative. However, by including 2:Independent, Intrade's mean efficiency increases to +0.8%, ranging from -6.5% to +6.3%. Instead of backing all possible outcomes, prediction markets allow an

Table 2: Comparing Bookmakers and Prediction Markets

		Primaries			Experiment		Runoff		
	Provider	V/E	I	ROI	I	ROI	V/E	I	ROI
Bookmakers	10Bet[1]						10.5%	never best	
	888sport				27.8%	5.0%	6.6%	12.7%	5.1%
	Bet365	7.3%	7.7%	5.8%	3.7%	4.9%	6.6%	0.9%	4.3%
	Betfred[1]						6.2%[2]	0.0%	7.9%
	Betsson[1]						5.3%[2]	0.0%	5.4%
	BlueSquare[1]	6.9%	3.4%	4.7%			6.4%	12.8%	5.1%
	Bodog		4.7%	3.7%			7.1%[2]	0.6%	9.0%
	Boylesports						7.3%[2]	3.8%	7.6%
	Bwin[1]						6.8%	0.0%	6.9%
	Centrebet[1]						5.9%	0.0%	6.1%
	Coral[1]						6.7%	0.0%	5.6%
	Intertops	8.1%	1.5%	9.7%	14.3%	4.9%	6.8%	6.4%	5.6%
	Ladbrokes		3.2%	5.8%	18.3%	6.5%	5.7%[2]	7.3%	9.9%
	PaddyPower	7.0%	17.8%	4.3%			6.0%	7.1%	6.0%
	PinnacleSports	2.2%	4.7%	4.1%	3.8%	4.4%	2.2%	0.2%	5.4%
	SkyBet	7.7%	2.3%	4.4%			6.6%	9.5%	5.6%
	Sportingbet	7.7%	4.9%	4.0%			7.5%	14.1%	5.7%
	StanJames	7.5%	7.1%	5.7%			6.3%[2]	5.1%	7.5%
	Totesport						5.8%[2]	5.1%	6.4%
	Unibet[1]						7.8%	0.0%	7.9%
	VictorChandler	8.6%	39.7%	5.5%			8.6%	2.2%	6.5%
	WilliamHill	6.9%	3.1%	4.4%	32.0%	5.3%	6.3%	12.2%	5.5%
Exchanges	Betfair	4.3%	22.8%	3.6%			4.4%	0.1%	42.7%
	Intrade	-1.2%	66.6%	5.9%	100%	4.6%	0.6%	97.9%	6.1%
	WBX	1.9%	10.6%	4.6%			2.3%	2.0%	11.5%

Notes: This table compares bookmakers and prediction markets according to their: (V/E) average vigorish or mean efficiency; (I) importance for surebets; (ROI) average return on investment of the best available surebets in the market. The meaning of the superscripts is as follows:

[1] Manual data extraction (every 8 hours)

[2] Average vigorish of Event 1 instead of Event 2

investor to lay on two outcomes in order to guarantee that at least one of them will occur. Intrade's mean efficiency by following this strategy is +2.1%, ranging from -0.6% to +16.0%.

In summary, bookmakers, as well as Betfair and WBX on their own, hardly ever offer surebet opportunities. This finding reflects a broad agreement in the literature that betting platforms, considered individually, are efficient (Vlastakis *et al.*, 2009; Tetlock, 2004; Gil and Levitt, 2007). The bookmakers' vigorish is typically higher than the exchanges' efficiency value. In addition, surebet opportunities exist on Intrade either by buying contracts on all three possible outcomes or by selling two of them. This finding is in line with insights from Rothschild and Sethi [2014]. In their paper, Rothschild and Sethi [2014] identify a handful of high-frequency traders who systematically exploit weak-form market inefficiencies on Intrade.

4.4 Intra-Market Surebet Opportunities

By construction, the use of different bookmakers or several exchanges provides surebet opportunities that are at least as profitable and as frequent as those conceived for a single platform. Such intra-market surebet opportunities among bookmakers have already been analyzed by Franck *et al.* [2013] and Vlastakis *et al.* [2009]. They find opportunities for intra-market surebets in less than 1% of the soccer games in their data.

Comparing all bookmakers' odds during the primaries reveals a mean vigorish of 2.1%, when considering only the outcomes 2:Democrats and 2:Republicans. If all possible outcomes are included, then the value increases to 3.2%. This intra-market vigorish is only slightly lower than the PinnacleSports' value and implies that PinnacleSports offers attractive odds on both outcomes. Intra-market surebets across bookmakers are possible in one-sixth of all observation points.

The average intra-market efficiency across exchanges is -2.2% when betting on 2:Democrats and 2:Republicans, -0.3% when backing on all outcomes, and -0.9% when laying bets on two outcomes. As mentioned in Section 4.2, these values include a 5% commission on Betfair and WBX. Even though surebet opportunities permanently exist, their mean return on investment is moderate. Across prediction markets, surebets can also be constructed by backing an outcome at one exchange and selling a contract on the same outcome at another prediction market. Betfair and WBX allow for only 5 surebet opportunities out of over 1.3 million intra-market comparisons on the outcomes 1:Obama, 2:Romney, 2:Democrats and 2:Republicans. The prices on these two platforms co-move very closely, which supports the finding of Wolfers and Zitzewitz [2004] that arbitrage opportunities between Tradesports and WorldSportsExchange are virtually absent due to the close co-movement of their prices. Rothschild and Pennock [2014] describe a persistent price spread between Intrade and Betfair which entails the opportunity for net earnings between 1 and 5%. My observations during the primaries confirm this spread. Positive return on investments, significant to the 1% level, are possible on 1:Romney by laying on Intrade and backing on either Betfair or WBX. The same is true for surebets by taking a short position on Intrade and a long position on Betfair for the outcomes 1:Obama and 2:Democrats. Even though the average ROIs are under 2%, surebets with an ROI of over 46% are occasionally possible.

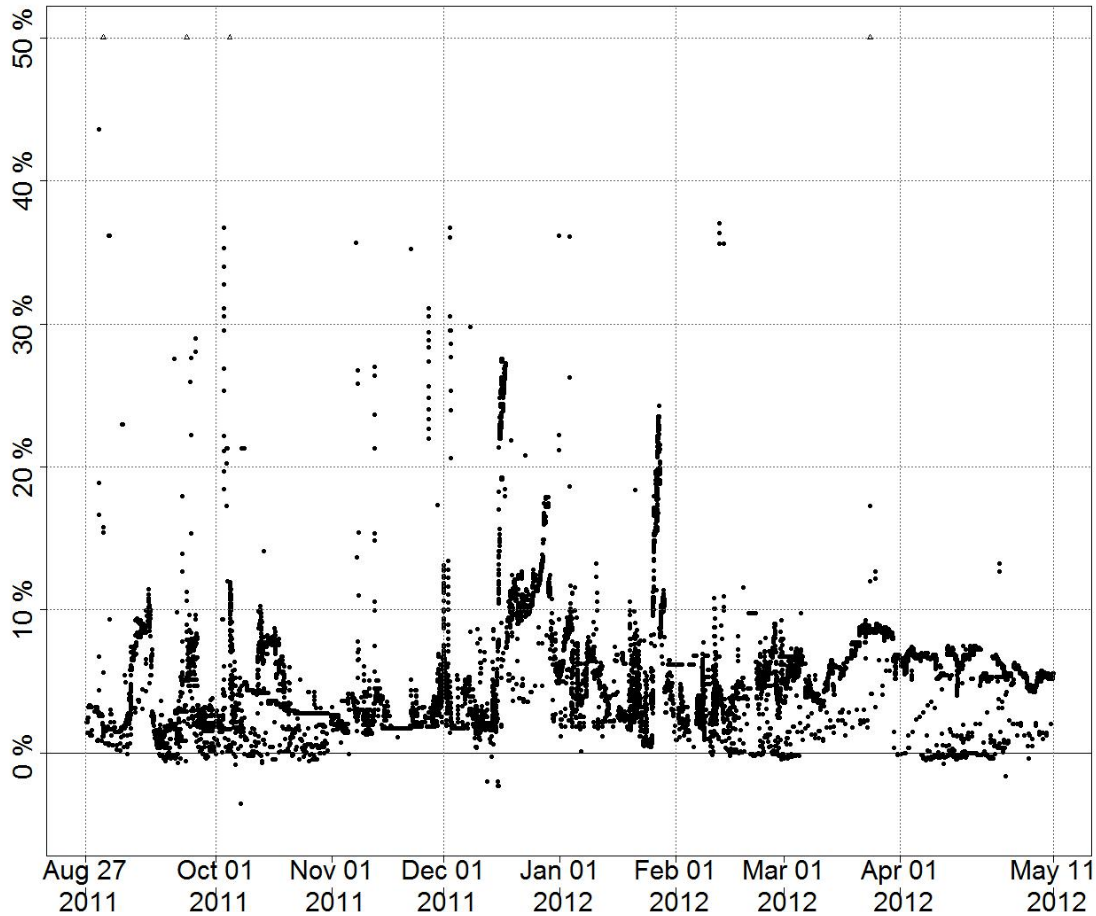
4.5 Inter-Market Surebet Opportunities

The most promising intra-market surebet construction investigated so far exploits price spreads across prediction markets on the same underlying outcome. However long positions can also be taken against bookmakers. In this subsection, I analyze short-position, inter-market surebet opportunities. Such surebet opportunities may exist on each potential outcome that is offered by at least one prediction market and one bookmaker. To identify short-position, inter-market surebets, I use the exchanges' lowest ρ_l and the bookmakers' highest ρ_b to calculate the return on investment. A positive ROI implies a surebet opportunity.

For each minute during the primaries, I calculated the highest available surebet opportunity over all outcomes and providers. Figure 1 shows the development of the best available ROIs.

The existence of inter-market surebet opportunities is not a new finding (Vlastakis *et al.*, 2009; Franck *et al.*, 2013). However, the permanent existence of surebet opportunities, with a mean of 5.3%, solely with bets on the 2012 U.S. Presidential Election is

Figure 1: ROI of the Best Available Surebets During the Primaries



Notes: This figure shows the return on investment of the best available surebets over all events, outcomes and providers during the primaries. Inter-market surebet opportunities (positive ROI) existed during 98.7% of all minutes. The mean ROI is 5.3% (std: 0.045) and the maximal ROI of an observed surebet reaches 311.79%. Surebets with a ROI greater than 50% are represented by a triangle.

remarkable.

In Figure 1, extreme values with a ROI over 50% are represented by a triangle. Curious offers from some individuals on Betfair are the cause of these inter-market surebet opportunities with a maximal ROI of over 300%. Within a couple of minutes these distortions were eliminated. I assume that the respective investors had placed these offers by mistake. These mistakes are caused by the different ways in which the providers express the offered odds. Another possible explanation for such extraordinary offers is that the investor either made a speculative attack or tested the market for efficiency, as Rhode and Strumpf [2006] did during the 2000 U.S. Presidential Election.

Before I address the two requirements and the hypothesis under investigation, I subdivide the best available ROIs by the respective underling outcomes and providers. The analysis of these best available ROIs aims to identify whether or not a focus on several providers and markets still allows most of the available best surebets to be identified. Because there are, at several points in time, either multiple outcomes and/or providers which result in the same surebet's ROI, the aggregate share of the respective underling outcomes or of providers, respectively, is typically larger than one. Thus, I define the index 'importance' as the share of the individual elements divided by the aggregate share of all the elements.

A total of 41 outcomes reflect the best ROIs, although 21 of these have an importance of less than 1%. Table 3 shows several indicators of the six most important and of four selected outcomes.

Table 3: Important Outcomes Reflecting the Best ROIs

Outcome	Importance	(1)	(2)	(3)	(4)	(5)	(6)
1:Romney	29.32%	5.9%	2.1%	10.2%	108,666	235,620	366,841
4:Romney	21.43%	7.9%	1.1%	149.0%	79,441	198,845	366,387
2:Republican	12.30%	4.4%	1.1%	13.2%	45,591	207,069	366,474
3:Warner	11.93%	3.0%	0.5%	21.3%	44,204	141,309	351,184
4:Gingrich	4.27%	3.9%	-1.0%	18.4%	15,818	114,720	366,485
4:Palin	4.07%	2.3%	-0.4%	18.0%	15,084	45,811	123,979
2:Democrat	0.64%	3.9%	-2.3%	5.8%	2,363	52,084	366,609
4:Santorum	0.26%	9.3%	-0.5%	36.8%	971	65,734	312,885
1:Obama	0.24%	2.0%	-3.1%	4.8%	874	14,379	366,862
1:Paul	0.18%	19.3%	-0.2%	180.0%	666	101,353	366,591
4:Perry	0.17%	5.9%	-2.3%	43.6%	632	28,479	205,149

Notes: This table shows the six most important and five selected outcomes reflecting the best ROIs. The additional indicators in this table describe the: (1) mean ROI if the surebet is the best in the market; (2) mean ROI; (3) maximum ROI; (4) number of best ROIs observed; (5) number of surebets observed; (6) total number of observations.

The six most important outcomes reflect more than 80% of all the best ROIs. The outcomes concerning Mitt Romney and the closely related outcome 2:Republican dominate the best ROIs. Column (2) shows that these outcomes also promise the highest mean return on investment. This relationship is self-evident, because ‘high importance’ implies that this outcome offered the best ROI in many observations. The values in column (1), the mean ROI, if the surebet is the best in the market, is susceptible to outliers, which are captured by the maximum ROI in column (3). Thus, the value of 19.3% for 1:Paul in column (1) is a consequence of the large outliers in column (3) and the low number of best ROIs observed in column (4). Column (6) presents the total number of observations for each event, given that at a specific point in time a lay and a back position was available. The majority of these values are close to the theoretically possible 371,520 observations and, thus, support the observation accuracy of the recording algorithm. On the one hand, missing observations may arise from changes in a website’s source code or owing to technical problems in the execution of the extraction algorithm; on the other hand, they may result from periodic website maintenance or may simply indicate that no odds were offered. The low observation values for 4:Palin and 4:Perry result from the fact that Sarah Palin and Rick Perry publicly stated that they suspended their candidatures on October 05, 2011, and January 19, 2012. In contrast, the lower value of 4:Santorum results from an error in the algorithm, which did not record Rick Santorum’s odds during October 2011.

Each of the three most important outcomes provide a surebet opportunity in more than two thirds of all observations, implied by the ratio of the values between column (5) and column (6).

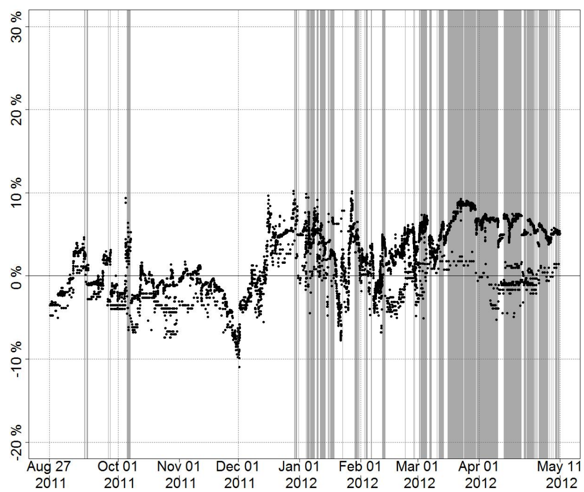
Figure 2 provides insights into the development of the odds and surebet opportunities for six important outcomes. These figures add a time dimension to the values in Table 3. The gray background indicates, that at the indicated time the outcome offered the market’s best surebet. The outcomes’ ROI develop a momentum of potential surebets or best ROIs in the market. Since the beginning of 2012, the outcomes for 1:Romney and 2:Republican are almost the only bets to offer the best ROI. Prior to that date, the outcomes for the primaries provided the most lucrative offers. The horizontal zero line separates potential arbitrage opportunities from sure losses. The corresponding figures on the right-hand side provide an insight into how the ROI fluctuations depend on the outcome's probability changes. The black points represent the best available lay position. The best available back position is represented by the upper boundary of the gray area. Of course, a continuous boundary is not completely correct and might imply the availability of a back position where none was available. However, because bookmakers adjust their odds less often than prediction markets, I justify using lines as this increases the readability of the figures. Surebet opportunities, or respectively a positive ROI, are available whenever the black dot lies above the gray area at a specific moment in time.

The odds of 1:Romney, 4:Romney and 2:Republican are related. The probability that Mitt Romney will win the presidential election equals the product of the probabilities that he will win the Republican primaries and that the Republican Party wins the presidential election. 4:Romney's Odds show three major drops that occur between November, 2011 and February, 2012. During these drops Mitt Romney was challenged by Newt Gingrich (first and second drop) as well as by Rick Santorum (third drop).

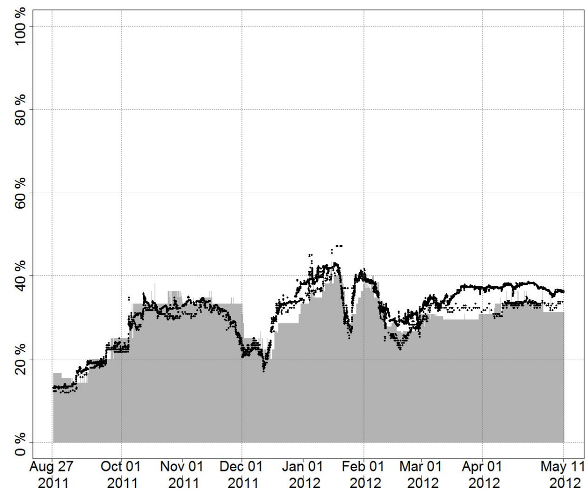
Prediction markets tend to react faster to new information than bookmakers, and this seems to be the reason for surebet opportunities whenever an outcome's probability rises. 3:Warner is a typical example, where the odds at a prediction market change, but the best bookmaker makes no adjustment to his odds. I use the term 'best bookmaker' to refer to the bookmaker who offers the currently lowest ρ_b in the market. The case of Rick Perry's lapse of memory in the Republican presidential primaries demonstrates the market's immediate reaction to such news. In the GOP debate on November 09, 2011, Rick Perry began an enthusiastic speech about his political program, promising to eliminate three government departments. Unfortunately, he was unable to remember all three of them and is quoted as admitting: "I can't. The third one, I can't. Sorry. Oops". The price of the 4:Perry stake immediately responded to that incident. Intrade's offered lay odds on his winning probability reacted after 3 minutes and dropped within 9 minutes from 8.7% to 3.0%. This immediate price reaction with following drift is in line with the findings of Gil and Levitt [2007] and Croxson and Reade [2014].

Figure 2: Surebet Opportunity and Odds Development of Six Important Outcomes

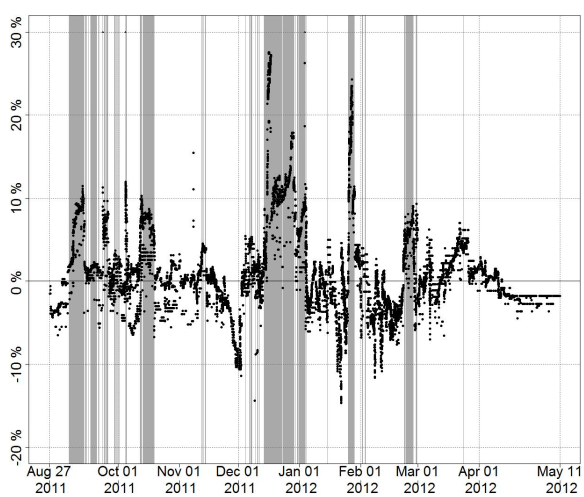
(a) 1:Romney ROI



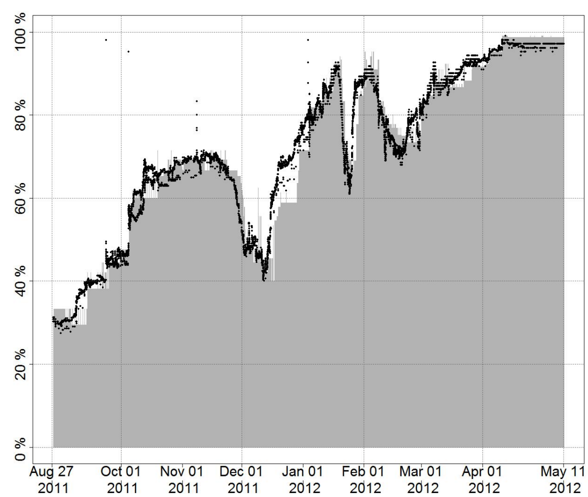
(b) 1:Romney Odds



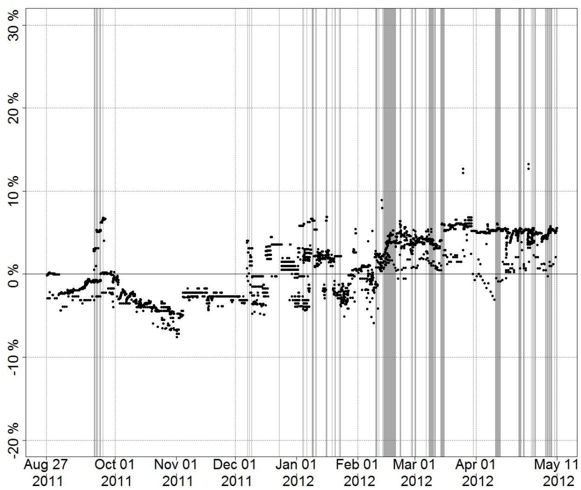
(c) 4:Romney ROI



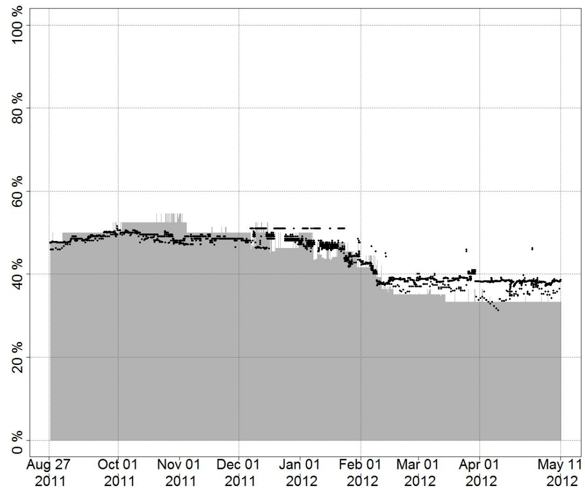
(d) 4:Romney Odds



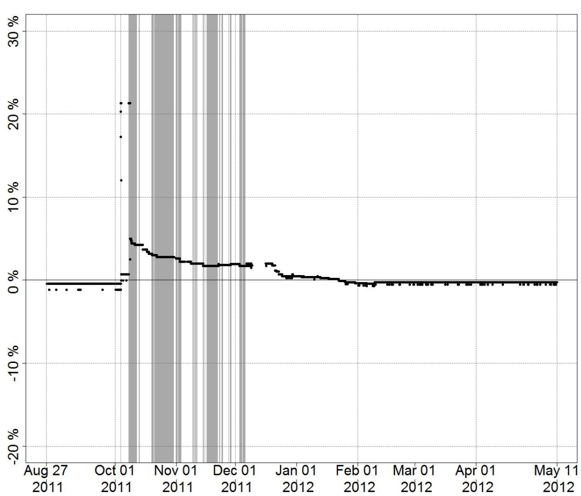
(e) 2:Republican ROI



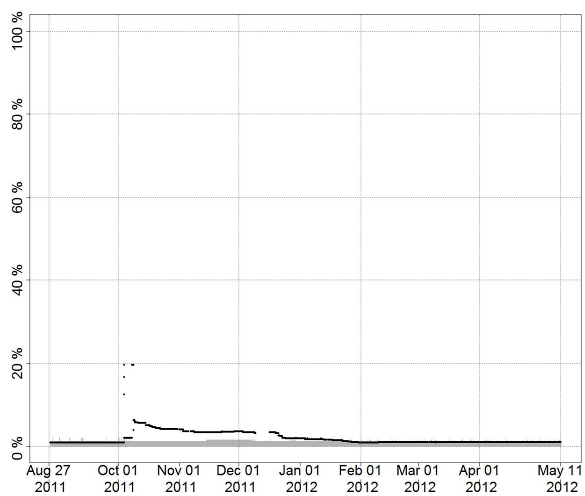
(f) 2:Republican Odds



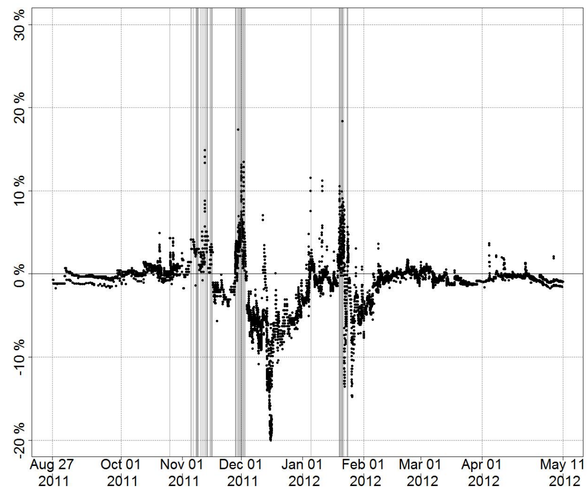
(g) 3:Warner ROI



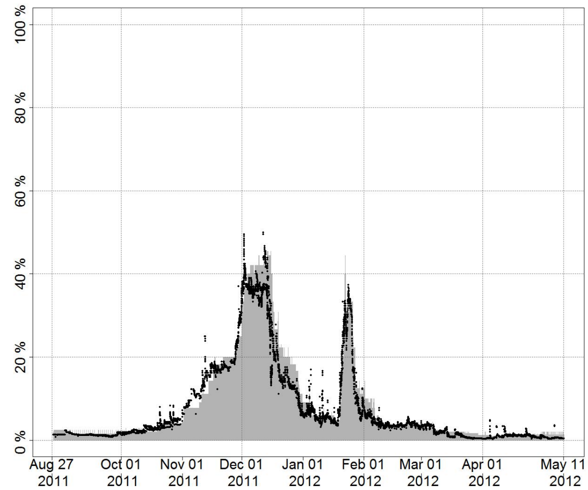
(h) 3:Warner Odds



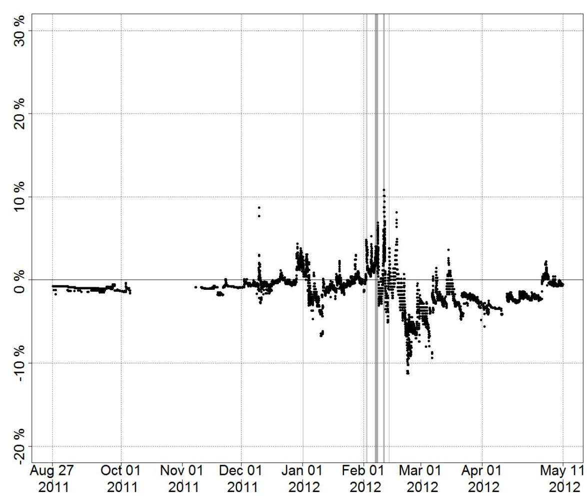
(i) 4:Gingrich ROI



(j) 4:Gingrich Odds



(k) 4:Santorum ROI



(l) 4:Santorum Odds

Notes: These figures show the development of surebet opportunities as well as the implied chance of winning the underlying event during the primaries. The gray background indicates that at this time the outcome offered the market's best available surebet. Surebet opportunities or positive ROIs, respectively, are available whenever the black dot lies above the gray area at a specific moment in time.

The best lay and back position may be driven by a single platform and therefore a general statement concerning the speed of acquiring new information is dangerous. In addition, other possible explanations must be considered. For example, if positive-feedback traders are present, then the model of De Long *et al.* [1990] can explain overreaction to news. Bookmakers whose odds reflect the true probability more accurately might then be forced to adjust their odds in order to rebalance the book. I address this question in Section 6 with respect to the runoff data.

The second column of Table 2 shows the distribution of the best ROI with regard to the related providers. VictorChandler and Intrade are by far the most important providers with regard to surebets during the primaries.

4.6 First Results

Only a few bookmakers and Intrade are responsible for the best available surebets. Surebets are typically possible when ρ_l increases, but not when ρ_l decreases. At a first glance, this seems strange as surebets should occur whenever the markets are not balanced. The reason for this observation might be that, on the one hand, laying against a bookmaker is not possible, and, on the other hand, exchanges seem to be the price leaders. Surebets are typically possible for events where the candidate is a close follower of the leader (e.g., 1:Romney) or is thrust into the spotlight (e.g., 4:Gingrich). Within a single platform, surebets are hardly ever possible and solely Intrade — the only betting exchange which charges a fixed fee instead of a commission — allows for a strategy that provides, on average, a positive return on investments. The analysis provides evidence for intra-market surebets, but the highest average ROIs exploit inter-market surebet opportunities.

Requirement 1 is definitely fulfilled in the underling events regarding the 2012 U.S. Presidential Election. Requirement 2 is not yet fulfilled. Although the information on current surebets is available, I have not yet considered whether or not the corresponding betting platforms are accessible. I deal with this last condition in the next section. In addition, I will check the following three observations regarding the primaries during the runoff in order to verify their validity.

Observation 1. *Concerning all four markets, surebets existed for 98.7% of the observation dates. They reached a mean ROI level of 5.3%. A surebet opportunity on an outcome was typically available for several weeks.*

Observation 2. *Surebets typically occurred when ρ_l increases and ρ_b remains low.*

Observation 3. *VictorChandler and Intrade are by far the most important providers.*

5 Runoff: Field Experiment

The previous section concludes that persistent surebets exist in the market and that the information required to identify them is publicly available for ongoing bets. This section presents a field experiment that I performed to highlight the risk and cost structure as well as the market size regarding investments in surebets. Furthermore, I outline the background to one of over 6,300 individual investors observed by Rothschild and Sethi [2014]. My experiment shows that inter-market arbitrage does not need to be excessively costly or complicated and reveals that 'Trader D' in Rothschild and Sethi [2014] probably used a strategy similar to mine. The most important objective of this study was to validate or reject my hypothesis that observed price inefficiencies across betting platforms can actually be exploited.

The most profitable and persistent surebets in the primaries occurred when the prediction markets reflected that an outsider had entered the market as a serious competitor. Rick Perry's 'Oops' moment demonstrated that candidates' statements in televised debates can influence prediction markets within a few minutes. I therefore precisely followed the activity of the online betting markets during the presidential debates.

5.1 Prediction Markets in the Field Experiment

I only used Intrade for laying bets on the exchanges. Betfair and WBX were not considered for the following two major reasons: Firstly, during the primaries, Intrade was responsible for most of the best available surebets in the market; Secondly, Betfair and WBX charge a 2.5%–5% commission on net winnings within a market. Even though the commission is not in itself a criterion for exclusion, the fixed monthly fee of USD 4.99 for using Intrade's services promises lower costs in comparison.

5.2 Bookmakers in the Field Experiment

The odds offered by 10Bet, 888sport, Betfred, Betsson, Boylesports, Bwin, Centrebet, Coral, Totesport, and Unibet were tracked in addition to the twelve bookmakers tracked for the primaries. Unfortunately, only 6 of the 22 bookmakers that were initially selected could be considered for the field experiment. These six bookmakers are: 888sport, Bet365, Intertops, Ladbrokes, PinnacleSports, and WilliamHill.

Bodog is a special case. On the one hand, I once tried to deposit money with Bodog, but my bank refused any transaction with it. On the other hand, Bodog closed its service on October 5, 2012. Even though Bodog was included at the beginning of the experiment, I excluded it from subsequent analyses. Sportingbet was not included, because their settlement rule differs from the other platforms.

Another fourteen bookmakers were not included, because they did not fulfill at least one of the following criteria: (1) Depositing money is free of charge; (2) Betting services are available to Swiss residents; (3) The user account can be administered in U.S. dollars; (4) No minimum turnover of the deposited money is required; or, (5) The information about the investor's individual limit is provided by the host prior to a deposit. While I required criteria (1) for reasons of simplification, criteria (2) was essential in order to fulfill Requirement 2. Finally, criteria (3), (4), and (5) provide essential information relating to arbitrage, because these restrictions exclude sources of currency risk as well as uncertain investment opportunities. Table 1 provides information about the reasons for excluding specific bookmakers.

All actions were performed in accordance with the platforms' terms and conditions. I disagree with Franck *et al.* [2013, p. 311] who assert in their paper that all bookmakers considered in their paper can cancel bets ex post, for example, on suspicion of arbitrage betting. In particular, their passage quoting WilliamHill's terms and conditions is misleading. The actual wording of the rules in the WilliamHill's [2012] terms and conditions, paragraph 11.4, are formulated as follows:

> If: [...] You have placed bets and/or played online games with any other online provider of gambling services and are suspected (as a result of such play) of any Prohibited Practice or otherwise improper activity [...] then, (including in connection with any suspension and/or termination of Your Account) we shall have the right, in respect of Your Account (and/or any other account held by You with a William Hill group company) to withhold the whole or part of the balance and/or recover from the account the amount of any deposits, pay-outs, bonuses or winnings which have been affected by or are in any way attributable to any of the event(s) contemplated in this paragraph 11.4. [...]

Using the services of different gambling websites is only forbidden *in connection* with fraudulent behavior; for example, cheating or collusion. Placing bets on the same event at different platforms itself is not prohibited. Upon request, WilliamHill affirmed to me that this is the correct interpretation of paragraph 11.4 in their terms and conditions.

5.3 Course of Action and Investments

The algorithm compared the current odds at least three times a day (around 1 AM, 9 AM and 5 PM EST) and sent me a market overview. If the algorithm reported any surebets exceeding a previously defined threshold, I exploited these. In the beginning of the experiment, I set the threshold at 5%, which is slightly lower than the average ROI during the primaries. During the experiment I increased this threshold in order to prevent myself from spending the available funds too early in the runoff. I set myself other terms such as a minimum U.S. dollar profit on the first bet with a new bookmaker or an obligation to split investments that exceed USD 900 in order to reduce the risk of unexpected circumstances.

Finally, I invested a total of USD 25,000 in 39 surebets on 1:Romney and 2:Republicans. The mean ROI of these surebets is 6.8%, with a maximum of 14.05%. The proceeds from the surebets amounting to USD 1,764 are offset by expenses incurred by service and transaction charges amounting to USD 134. Appendix A.1 lists all investments of this field experiment.

I started the field experiment with an investment capital of around USD 11,000. In

July 2012, see Figure 3(b), the ROI on available surebets decreased, and I thought that the market would become more efficient as the election approached. Owing to my — in retrospect mistaken — assessment, I invested my remaining funds up until August 2, 2012. A few days later, the ROI improved substantially, and new funds allowed me to continue betting on surebets.

The money invested with the prediction market is frozen in the account until the settlement of the bet. Intrade accept credit card deposits, but only up to USD 2,000 over a 30-day period. Credit card payments have a major advantage, as transfers to the user accounts take only a few seconds and thus prevent idle money from accumulating on users' accounts. Bookmakers generally allow much higher credit card deposits than Intrade. As I laid bets solely on Intrade, I deposited larger amounts in advance in my Intrade account by means of international wire transfer in order to avoid this friction. These transfers were credited to my Intrade account the next working day.

Information about current odds were extracted and compared automatically. On the other hand, transferring money and placing bets were manual tasks. My investment order for the first 2 bets was: Place lay investments first and then invest in the back position. My idea was that it might be safer to first invest in the market that changed more frequently. In the second bet, I did not check the maximum stack size at 888sport in advance. I then had to use Bet365 to hedge the lay positions that I had already bought. Even though this was my mistake, I decided to change the investment order, because bookmakers are not committed to the odds they offer. A second mistake of mine, which I detected a few days later, occurred on August 2, 2012, when I bought too many contracts on WilliamHill.

After the bets were settled on November 7, 2012, I withdrew all funds from Intrade. Intrade dispatched my withdrawal request on November 14, 2012. Two working days later, the money was credited to my bank account.

5.4 Risks

Potential risks of the field experiment's investments are:

Currency risks I precluded currency risks by selecting bookmakers and exchanges which permit U.S. dollar accounts. In addition, all account fees and credit cards were administered in U.S. dollars.

Counterparty risks Three contractual relationships in the field experiment bear counterparty risks.

1. Bookmakers announce their odds as well as the (individual) stack limits. These offers are non-binding and the bookmaker can reject bets without explanation. But once they have accepted a request and assign a reference number to the bet, the contract is binding. The change in the investment order after my first two bets account for this uncertainty. I agree with Franck *et al.* [2013, p. 311] that a bookmaker can back out of a contract if the investor has violated the terms and conditions of the platform. Therefore, it was important to act in compliance with the provider's terms and conditions.
2. Since Intrade freezes the money to cover potential losses on the investor's account, the contracting partner faces no risk.
3. The operator of the platform is obliged to provide the operational services and manage the money accounts. The risk that the platform might delay withdrawals or, in the worst case, becomes bankrupt cannot be ruled out. This insolvency-risk must be kept in mind.

Unlikely circumstances This risk occurs if an unlikely circumstance prevents unambiguous settlement. In sports, it is possible for events to be cancelled or abandoned. However, as the markets under investigation refer to the 2012 U.S. Presidential Election, the threat of this risk is negligible.

5.5 Costs

Gains from surebets must cover their expenses in order to provide profitable investment opportunities. Such expenses result from:

Service costs While the accounts on bookmakers platforms are free of charge, Intrade charges a fixed monthly fee (USD 4.99) for its services. Intrade and VictorChandler charge a fee for withdrawals of USD 20 and USD 24, respectively, whereas other providers offer this service without charging a commission. Neither my bank account nor my pre-paid credit card carried charges.

Transactions These fees cover deposits to the credit card (USD 3) and commissions on international wire transfers either by my bank (at least USD 6) or by any intermediate banks (typically around USD 20).

Search costs Identifying surebets by checking current quotes manually would result in high opportunity costs. These variable costs present an argument against the existence of arbitrage. In the context of this paper, I wrote a program that automatically extracts the current odds, identifies surebets and periodically transmits this information to my smartphone. Therefore, my search costs were fixed, and the program could be adjusted for forthcoming events.

Cost of capital The invested capital was an interest-free loan, and thus the cost of capital was zero. At the current rate of interest, the cost of capital is generally negligible.

Taxes For Swiss residents, net profit on the applied investment strategy is subject to Federal Direct Tax according to Article 16 DBG. The actual tax due will be a fraction of the net gain (after all costs are subtracted), and will therefore not impact the profitability of the strategy.

Table 4: Experiment's Total Service and Transaction Costs

Intrade's fixed service fee for six months	USD 29.94
Transactions to the pre-paid credit card	USD 12.00
Wire transaction to Intrade (three deposits)	USD 23.00
Intrade's service fee for withdrawal	USD 20.00
Various intermediate-bank charges for wire transactions	USD 49.04
Total	USD 133.98

Table 4 shows the composition of the total service and transaction costs which accrued throughout the field experiment. As these costs are almost fixed, they will diminish as investments in surebets increase.

5.6 Market Size

Bookmakers' maximum betting limits determined the long side of the market. Each bookmaker's terms and conditions contain information about its maximum betting limit. But this value, which typically ranges between USD 200,000 and USD 1,000,000, is almost always disproportionate to the investor's individual limit. Ladbrokes and PinnacleSports stand out from the other bookmakers. Not only do they provide a considerably higher individual betting limit of over USD 7,000 per day, they also subsequently reset the limit back to this level within 24 hours. My available capital did not allow me to fully test these limits, but on August 24, 2012 (Bet 27) I invested USD 1,900 on PinnacleSports without any problem. WilliamHill and 888sport also increased the individual limit after bets were placed, but at a much lower level. On both platforms, this individual restriction limited investments to profitable surebet opportunities. 888sport allowed bets for a total of only USD 366, whereas bets of nearly USD 4,000 were possible on WilliamHill. Bet365 (USD 954) and Intertops (USD 821) do not increase the individual limit after a bet has been placed. Nevertheless, large surebet opportunities such as that on November 6, 2012 08:37 PM EST are possible. At that point in time, Ladbroke's limit of USD 8,050 allowed for an investment amounting to USD 53,895.36 at a ROI of over 19.5%, and repayment within 12 hours to be made. This amount is nowhere near the size of investments made by Wall-Street traders, but is quite large for most individual investors. Due to lack of information regarding the exchanges' full order book, bookmakers' limits, market accessibility and other restrictions, these analyses reveal only a lower bound of the actual market size for a Swiss resident.

Higher liquidity on exchanges may or may not increase the platform's efficiency and lower surebet opportunities. The decisive factor is who provides the liquidity. Arbitrageurs certainly reduce price differences within or across providers. However, Rothschild and Sethi [2014] find strong indications of a manipulator who provided liquidity in order to shield a high price on 1:Romney contracts. This price shield is also visible in Figure 4(b).

5.7 Results

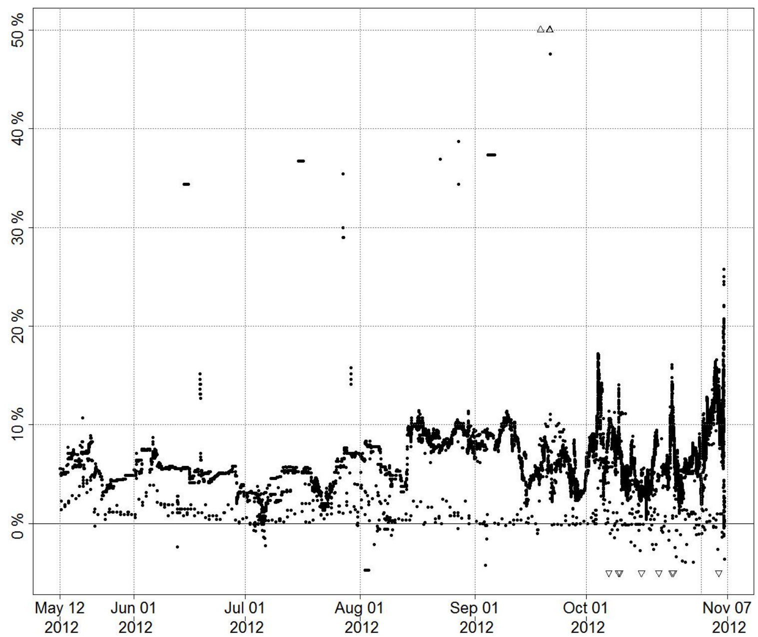
Having shown the existence of surebets in Section 4, I check Requirement 1 for the runoff. Figure 3(a) is a counterpart to Figure 1 and shows the best observed surebets at each point in time. Since surebets existed almost continuously during the runoff, Requirement 1 is clearly fulfilled and Observation 1 confirmed.

A surebet is exploitable if the necessary information on the surebet is available at the time of its occurrence and if the markets are accessible. The first condition is satisfied, as the odds were observed and analyzed in real time. For the purposes of the second condition, Figure 3(b) again shows the best observed surebets at each point in time. However, this time, only Intrade and the six bookmakers considered in field experiment are included in the analyses. The mean ROI thereby decreases from 6.3% to 5.5%, and the probability that at a random minute in time an exploitable surebet opportunity exists decreases from 96.7% to 95.7%. The bookmakers included in the experimental setting are all accessible to Swiss residents, thus fulfilling Requirement 2. I expected a falling ROI such that the annualized ROI would remain more or less stable. The opposite turned out to be the case: As the election day drew closer, the ROI became more volatile. This observation, in association with the insight of Rothschild and Sethi [2014] regarding the successful price shield of the manipulator, implies that biased traders, as well as noise traders, heavily dominated unbiased traders in the market in the last few weeks leading up to the election. With the benefit of hindsight, the most profitable tactic would have been to invest all the money on the election day of November 6, 2012 08:37 PM EST at a ROI of over 19.5%, repayment within 12 hours.

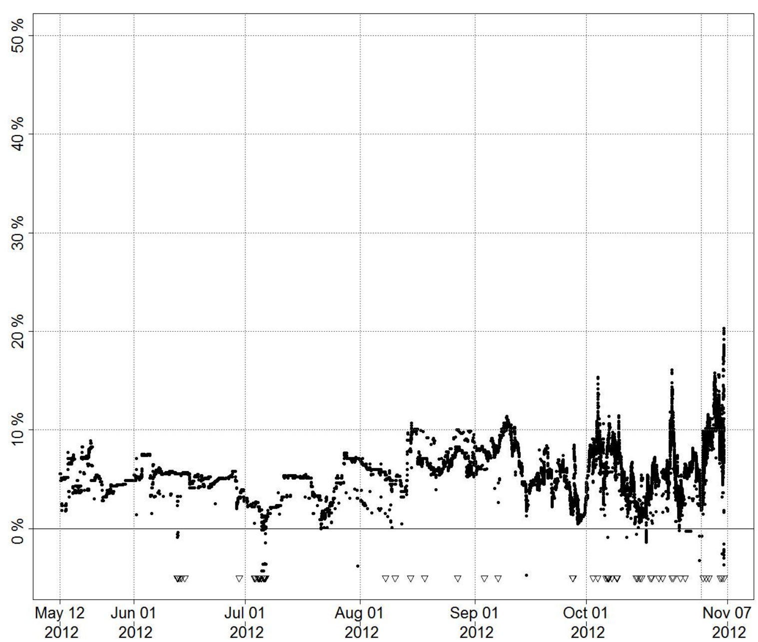
My experiment's arbitrage-based strategy resulted in the following characteristics used by Rothschild and Sethi [2014]. My strategy's value for what they call 'aggression' equals 1 and results from taking liquidity out of the market by matching offered lay positions. Duration = 1 and direction = +1 result because all bets on Intrade were short, either on 1:Romney or 2:Republicans, and were held to expiration. During the whole runoff,

Figure 3: ROI of the Best Available Surebets During the Runoff

(a) Runoff



(b) Experiment



Notes: These figures show the return on investment of the best available surebet either over all events, outcomes and providers during the runoff, or the subset used for the experiment. ROIs greater than 50% or smaller than -5% are represented by a triangle.

my experiment contained 77 trades with a total volume of 2,772 contracts. These bets required an amount of USD 15,989.29 to be deposited as a security which was frozen on Intrade to compensate for potential losses. Barack Obama's re-election resulted in a profit of USD 11,730.71 on Intrade. These characteristics are close to the Trader D's values in Rothschild and Sethi [2014], even though this investor had a much larger wallet. The insights of my experiment open up a new perspective regarding his behavior, which was probably not risky at all. From my point of view, he used an inter-market, arbitrage-based strategy similar to mine. If my assumption is correct, then a large share of the profit that he gained on Intrade was used to compensate his unobserved losses on other providers.

Having discussed surebets in detail, I now turn to the last requirements for arbitrage. Whether or not the proceeds from the field experiment are able to cover all the costs accrued in the investigation depends on the dollar-equivalent amount incurred by the search costs, or, more precisely, the expenses incurred in developing the computer program. However, as already discussed in Subsection 5.6, the market would have allowed for substantially larger investments in exploitable surebets. The exploitable opportunity available on November 6, 2012, alone, would have covered all the costs incurred, and in itself confirms that market size permits profitable inter-market surebet strategies.

The surebet strategy employed in this field experiment in inter-market betting was developed with the objective of excluding any risk of losing money. Even though such riskless profits might not maximize an investor's utility function, the expected profit of risk carrying behavior can be compared to the gains of these exploitable surebets. The only remaining risk that I am not able to avoid and which, therefore, must be considered is the insolvency-risk. The importance of considering this risk is demonstrated by Intrade's near insolvency in 2013. A short summary on this incidence is given in Appendix A.2.

6 Runoff: Market Analyses

The vigorish values for the ten bookmakers, shown in Table 2, have been calculated for the primaries as well as for the runoff. These values are highly positively correlated (0.96) and are on average slightly lower during the runoff. 10Bet has the highest vigorish (10.5%)

of all bookmakers. Because such a high value indicates that 10Bet offers the worst odds for at least one outcome, it is not astonishing that 10Bet never offered the back position for the best available surebet in the market. PinnacleSports offers, as in the primaries, by far the lowest vigorish of all the bookmakers. However, PinnacleSports' importance is low. Even though PinnacleSports features only half as many observations compared to the most important platforms, the reason for the low level of their importance is that their low vigorish results from high odds on the outcome 2:Democrats. PinnacleSports is responsible for 80% of all surebets on 2:Democrats, but their importance is low, because these surebets do not offer the best ROI in the market. Although the platform's vigorish may be interpreted as indicating lucrative surebets, the odds' distribution is more indicative.

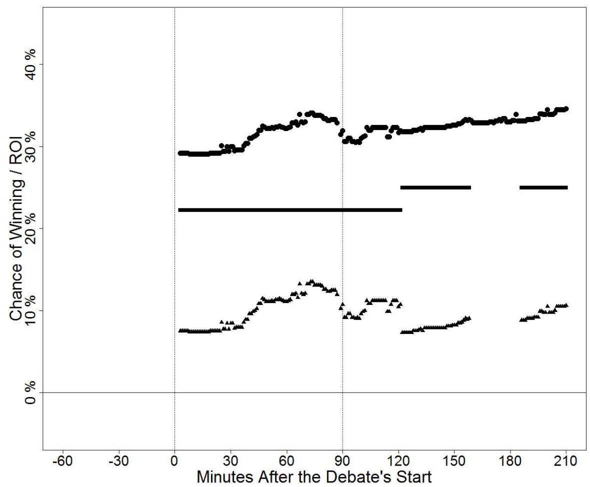
Intrade is responsible for almost all best surebets during the runoff. In addition, Intrade's average ROI of the best available surebets in the market increase compared with the primaries. The importance of the platforms vary considerably between the primaries, the experiment and the runoff. VictorChandler, the most important bookmaker during the primaries, is negligible throughout the runoff. In contrast, WilliamHill only plays a significant role during the months close to the election. These variations show that the selection of bookmakers for the assessment as well as for the observation period has a significant impact on the markets' analyses. Even though Intrade is still the most important prediction market, Observation 3 cannot be verified. The importance as well as the average ROI of the providers' best surebets vary strongly, such that I cannot make generalized statements about a bookmaker's importance.

Observation 2 identifies, that the slow reaction of some bookmakers to new information in the market is a typical cause of surebet opportunities. This is also true for some observations in the runoff and is also in line with the findings of Croxson and Reade [2011] on games played during the Euro 2008 soccer championship. Croxson and Reade's [2011] analyses show that traditional bookmakers (Ladbrokes and WilliamHill) follow the price discovery of the betting exchange (Betfair). The best examples of these surebets occurred during the (vice-)presidential debates. Figures 4(a)-(d) show that the odds on Intrade were volatile during these specific hours. Most bookmakers appear to have taken the impact of the debates on the public into account and closed their markets for a few hours, while other bookmakers offered odds without adjustments during the debates, and thereby gave rise to high surebet opportunities. During the first presidential debate (see Figure 4(a)), for example, WilliamHill was the only bookmaker in the field experiment who was still offering odds on 1:Romney at this point. Once the debate was over, WilliamHill adjusted the odds slightly. The jump in price of the best available back offer at the beginning of the second presidential debate occurred because WilliamHill (and Intertops) stopped offering odds on the presidential election, leaving Ladbrokes as the only bookmaker in the experiment. Event 2 reveals that Bet365 offered odds on 2:Republicans during all four debates.

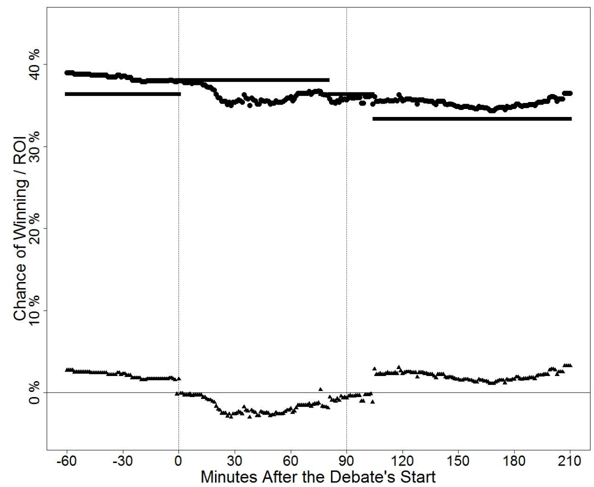
However, slow bookmakers were not the reason for the highest available surebets. On the contrary, in September 2012 and on election day some bookmakers seemed to anticipate the downward trend on 1:Romney and 2:Republicans. Bookmakers pushed ρ_b downwards several times, which resulted in surebet opportunities with a ROI of up to 10%, and Intrade's investors adjusted the ρ_l afterwards. During September 2012 (see Figure 5(a)), the driving force behind ρ_b were Ladbrokes and, later on, Sportingbet, which were together responsible for more than two thirds of the best odds' changes. 1:Romney underwent more changes than 1:Obama, because the best bookmaker's odds change whenever a single bookmaker offers odds that imply a lower chance of winning than any other bookmaker in the market. In contrast to this, the best bookmaker's odds on 1:Obama increases only when the last bookmaker adjusts his odds. In this case, Intertops is responsible for the first two price jumps, Totesport for the third, and SkyBet's odds adjustment on September 27, 2012, resulted in the forth jump. Figure 5(b) shows the development of the winning chances for 1:Obama and 1:Romney on the election day. At 7 PM EST the first polling stations were closed. The vertical line at 11:15 PM indicates the point in time when the media projected that Barack Obama would carry the swing-state Ohio. The last vertical line indicates the point in time at which Mitt Romney congratulates Barack Obama on his successful re-election in his concession

Figure 4: Information Processing: (Vice-)Presidential Debates

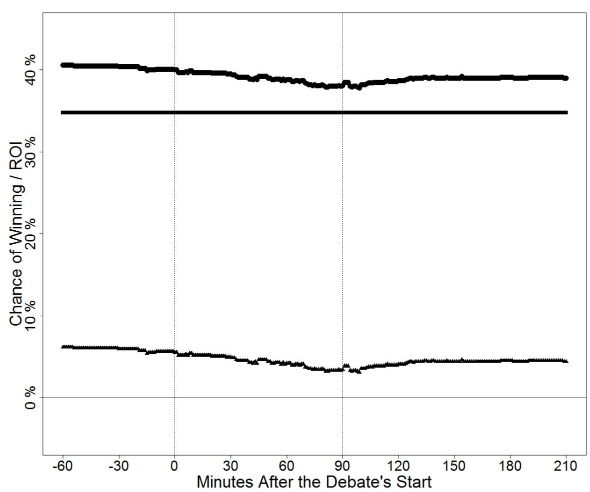
(a) Obama vs. Romney on October 04, 2012



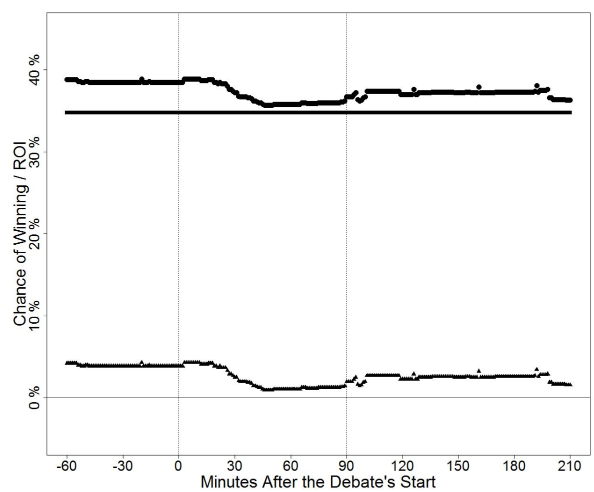
(b) Obama vs. Romney on October 17, 2012



(c) Obama vs. Romney on October 23, 2012



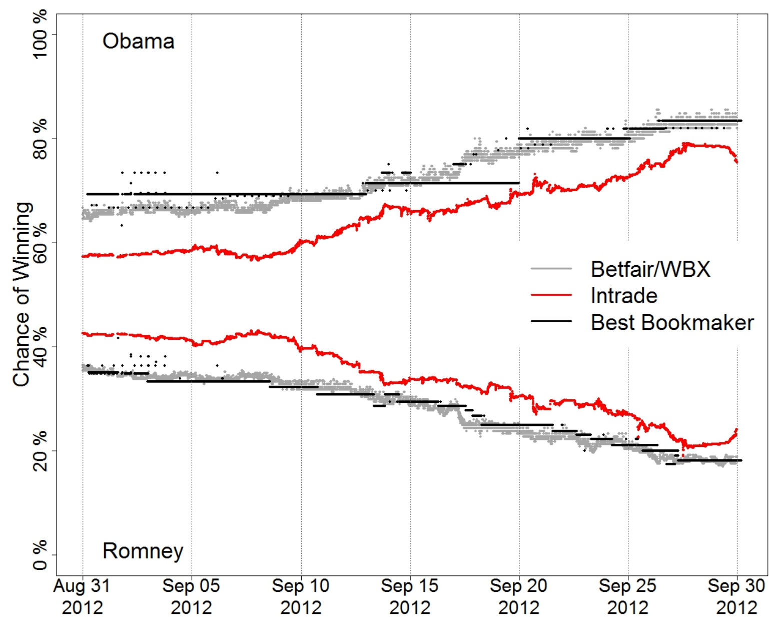
(d) Biden vs. Ryan on October 12, 2012



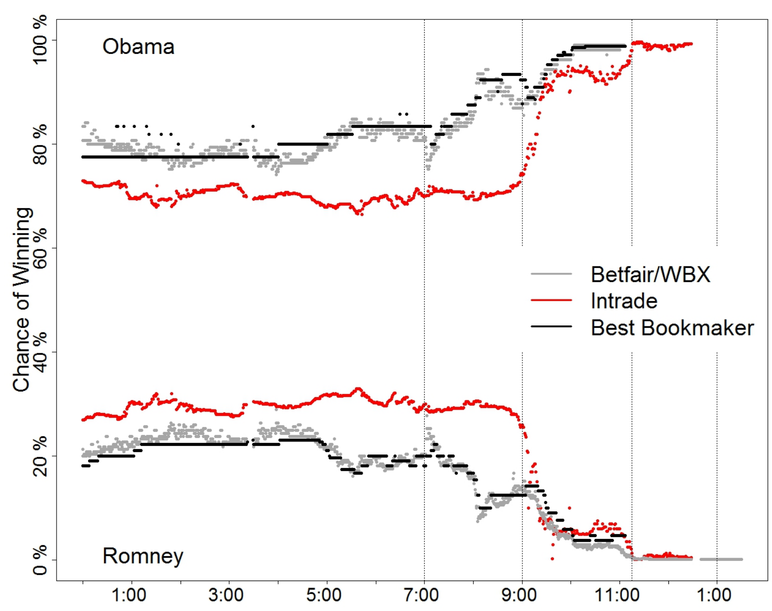
Notes: This figure shows examples of the information processing of prediction markets and bookmakers during the (vice-)presidential debates. These televised debates lasted for 90 minutes. The upper volatile line represents ρ_l; the mostly stable course is ρ_b; and the lower volatile line represents the current surebet's ROI.

Figure 5: Information Processing: September 2012 and Election Night

(a) September 2012 Odds Development



(b) Election Day's Odds Development



Notes: This figure shows examples of the information processing of prediction markets and bookmakers in September 2012 and during the election night.

speech. Although the prices on Intrade implied that 1:Romney remained stable until 9 PM, ρ_b as well as the price for open lay positions on Betfair and WBX already began to decrease from 5 PM onwards. Rothschild and Sethi [2014] show in their analyses that the price on Intrade was shielded by a single large investor who prevented a price drop. Before this point in time, Ladbrokes offered the best quotes. After this point, StanJames, together with Sportingbet and PaddyPower, took the lead. StanJames' quotes were responsible for the changes in ρ_b until 10 PM, when StanJames closed its markets.

7 Conclusion

In this paper, I provide evidence that arbitrage opportunities in inter-market online betting exist and can be exploited. I wrote an algorithm which observed the minute-by-minute odds on the 2012 U.S. Presidential Election for 4 prediction markets and a total of 22 bookmakers. The primaries served as a trial period which enabled me to confirm that the most important requirements concerning arbitrage opportunities are fulfilled. As a field experiment, I acted as an arbitrageur in the market during the runoff to the election and invested over USD 25,000 in surebets. Besides establishing a sure mean ROI of 6.8% on these bets, I identified the costs and risks of this arbitrage-based strategy. On balance, a sure profit of USD 1,630 resulted. The strategy's set-up was such that all risks, except the insolvency-risk, were eliminated. The total costs were almost all fixed costs, and thus became increasingly smaller as the amount of investment increased with available surebet opportunities.

The market size and the ROIs of the surebet opportunities that were available during these events outweighed the remaining risks and personal opportunity costs incurred in writing the required algorithm. I conclude that my hypothesis is verified, and that this market offered arbitrage opportunities. In addition, I show that arbitrage opportunities persistently exist and that they can be found and exploited with moderate effort. A crucial condition for this finding to hold is that these sure and riskless profits need an inter-market approach which, includes bookmakers as well as prediction markets. Even though I show that riskless profits are exploitable, I do not state that this strategy is optimal and that investors should actually avoid risk at all costs.

Media attention, new information which becomes continuously available, potential influence of the odds on voting behavior, and, for example, the number of upcoming events are all factors that differentiate political betting from sports betting. Therefore, the results' application regarding the persistence and the level of available arbitrage opportunities on sports events remains uncertain. In addition, I judge market accessibility and, for example, the transaction costs only from the perspective of a Swiss resident. However, regarding the cost structure of adopting an arbitrage-based strategy this paper provides concrete and applicable values.

In contrast to the existing literature, where the authors only use data on past events, I not only show that sufficient information on current arbitrage opportunities is available at negligible variable search costs, but also identify the accessibility of the involved markets. This becomes important with regard to the fact that only 6 out of 22 observed bookmakers satisfied the requirements for placing bets with them according to my suggested arbitrage-based strategy.

Even though I do not use the odds in order to predict the event's outcome, I highlight that the odds' flexibility on one betting platform might be restricted because of the prices offered on other platforms and, therefore, by the law of one price. For example, the contract prices on Intrade might not fully absorb new information, because the odds on other platforms are conditioned by price adjustments caused by arbitrageurs.

Finally, my suggested arbitrage-based strategy offers new insights into the possible objectives that some investors may have, in view of the puzzling behavior that they exhibit when only observing the investments on one platform. For example, the strategy of one individual (Trader D) analyzed by Rothschild and Sethi [2014] seems to be very risky at first sight. However, this individual has probably placed investments according to my strategy and, thereby, exploited inter-market arbitrage opportunities.

It would be interesting to investigate whether or not my suggested strategy is generally applicable to sports betting. Further research on identifying information leaders within the market and combining information from bookmakers' and exchanges' quotes with opinion polls might enhance the discussion on forecasting the underlying event's outcome. However, this paper demonstrates that it is worth taking a step back in order to look at the online betting market as a whole, rather than focusing on a single provider.

A Appendix

A.1 Detailed Trading Information

Date/Bet Nr.	LayProvider	LaySize	LayChance	LayInvest	LayTotalInvest	BackProvider	BackInvest	BackChance	BackTotalInvest	Profit Rep. Win	ROI Rep. Win	Profit Rep. Lose	ROI Rep. Lose
Surebets on 2:Republicans													
05/12	Intr	36	38.0%	223.2		Inter	120	33.3%		16.80	4.90%	16.80	4.90%
1					**223.2**				**120**	**16.8**	**4.90%**	**16.8**	**4.90%**
06/05	Intr	1	45.1%	5.49		888	3.82	38.2%		0.70	7.50%	0.69	7.41%
		12	45.1%	65.88		888	45.8	38.2%		8.32	7.45%	8.32	7.45%
		35	45.0%	192.5		888	133.58	38.2%		23.90	7.33%	23.92	7.34%
						888	1.8	38.2%		2.92	162.00%	-1.80	-100.00%
		46	45.0%	253		B365	173.32	42.2%		-15.55	-3.65%	33.68	7.90%
		2	44.0%	11.2		B365	7.68	42.2%		-0.68	-3.59%	1.12	5.93%
		1	43.1%	5.69						-5.69	-100.00%	4.31	75.75%
2					**533.76**				**366**	**13.91**	**1.55%**	**70.24**	**7.81%**
06/07	Intr	1	45.7%	5.43		888	4	40.0%		0.57	6.04%	0.57	6.04%
		5	45.7%	27.15		888	20	40.0%		2.85	6.04%	2.85	6.04%
3					**32.58**				**24**	**3.42**	**6.04%**	**3.42**	**6.04%**
06/15	Intr	5	45.4%	27.3		888	20	40.0%		2.70	5.71%	2.70	5.71%
		3	45.4%	16.38		888	12	40.0%		1.62	5.71%	1.62	5.71%
		2	45.3%	10.94		888	8	40.0%	40	1.06	5.60%	1.06	5.60%
		6	45.3%	32.82		Inter	24	40.0%		3.18	5.60%	3.18	5.60%
		101	45.2%	553.48		Inter	404	40.0%	428	52.52	5.49%	52.52	5.49%
4					**640.92**				**468**	**61.08**	**5.51%**	**61.08**	**5.51%**
06/18	Intr	2	45.2%	10.96		888	8	40.0%		1.04	5.49%	1.04	5.49%
		1	45.0%	5.5		888	4	40.0%		0.50	5.26%	0.50	5.26%
		1	45.0%	5.5		888	4	40.0%		0.50	5.26%	0.50	5.26%
5					**21.96**				**16**	**2.04**	**5.37%**	**2.04**	**5.37%**
06/19	Intr	3	45.0%	16.5		888	12	40.0%		1.50	5.26%	1.50	5.26%
6					**16.5**				**12**	**1.50**	**5.26%**	**1.50**	**5.26%**
06/26	Intr	5	44.9%	27.55		888	20	40.0%		2.45	5.15%	2.45	5.15%
		1	44.9%	5.51		888	4	40.0%		0.49	5.15%	0.49	5.15%
		4	44.8%	22.08		888	16	40.0%		1.92	5.04%	1.92	5.04%
7					**55.14**				**40**	**4.86**	**5.11%**	**4.86**	**5.11%**
07/11	Intr	5	43.2%	28.4		WHill	19.05	38.1%		2.56	5.39%	2.55	5.37%
		9	43.1%	51.21		WHill	34.25	38.1%		4.45	5.20%	4.54	5.31%
8					**79.61**				**53.3**	**7.00**	**5.27%**	**7.09**	**5.33%**
07/11	Intr	5	43.1%	28.45		WHill	19.05	38.1%		2.51	5.28%	2.50	5.26%
		130	43.0%	741		WHill	495.15	38.1%		63.62	5.15%	63.85	5.17%
9					**769.45**				**514.2**	**66.13**	**5.15%**	**66.35**	**5.17%**
07/12	Intr	120	43.1%	682.8		WHill	457.14	38.1%		60.05	5.27%	60.06	5.27%
		74	43.0%	421.8		WHill	281.86	38.1%		36.22	5.15%	36.34	5.16%
10					**1104.6**				**739**	**96.27**	**5.22%**	**96.40**	**5.23%**
07/13	Intr	10	43.2%	56.8		WHill	38.1	38.1%		5.11	5.39%	5.10	5.37%
11					**56.8**				**38.1**	**5.11**	**5.39%**	**5.10**	**5.37%**
07/16	Intr	1	43.1%	5.69		WHill	3.81	38.1%		0.50	5.28%	0.50	5.26%
		62	43.0%	353.4		WHill	235.83	38.1%		29.82	5.06%	30.77	5.22%
12					**359.09**				**239.64**	**30.33**	**5.06%**	**31.27**	**5.22%**
07/17	Intr	24	43.1%	136.56		WHill	91.43	38.1%		12.01	5.27%	12.01	5.27%
		208	43.0%	1185.6		WHill	792.37	38.1%		102.00	5.16%	102.03	5.16%
13					**1322.16**				**883.8**	**114.02**	**5.17%**	**114.04**	**5.17%**
07/18	Intr	15	43.0%	85.5		WHill	57.14	38.1%		7.35	5.15%	7.36	5.16%
		5	43.0%	28.5		888	20	40.0%		1.50	3.09%	1.50	3.09%

14					**114**				**77.14**	**8.85**	**4.63%**	**8.86**	**4.64%**
07/19	Intr	1	41.5%	5.85		WHill	3.81	38.1%		0.34	3.53%	0.34	3.52%
		2	41.4%	11.72		WHill	7.62	38.1%		0.66	3.43%	0.66	3.41%
		12	41.3%	70.44		WHill	45.71	38.1%		3.84	3.30%	3.85	3.31%
15					**88.01**				**57.14**	**4.84**	**3.34%**	**4.85**	**3.34%**
07/20	Intr	5	41.3%	29.35		WHill	19.05	38.1%		1.61	3.32%	1.60	3.31%
		6	41.2%	35.28		WHill	22.86	38.1%		1.87	3.21%	1.86	3.20%
		4	40.9%	23.64		WHill	15.24	38.1%		1.13	2.89%	1.12	2.88%
16					**88.27**				**57.15**	**4.60**	**3.16%**	**4.58**	**3.15%**
07/22	Intr	1	41.2%	5.88		WHill	3.81	38.1%		0.31	3.21%	0.31	3.20%
		14	41.1%	82.46		WHill	53.33	38.1%		4.20	3.09%	4.21	3.10%
17					**88.34**				**57.14**	**4.51**	**3.10%**	**4.52**	**3.11%**
07/24	Intr	5	41.1%	29.45		888	20	40.0%		0.55	1.11%	0.55	1.11%
18					**29.45**				**20**	**0.55**	**1.11%**	**0.55**	**1.11%**
08/02	Intr	2.25	42.5%	12.9375		888	9	40.0%		0.56	2.56%	0.56	2.56%
		6.75	42.5%	38.8125		WHill	24.55	36.4%		4.15	6.55%	4.14	6.53%
						WHill	6.38	36.4%		11.17	175.00%	-6.38	-100.00%
19					**51.75**				**39.93**	**15.88**	**17.32%**	**-1.68**	**-1.83%**
08/02	Intr	8	42.5%	46		B365	32	40.0%		2.00	2.56%	2.00	2.56%
20					**46**				**32**	**2.00**	**2.56%**	**2.00**	**2.56%**
08/17	Intr	2	43.0%	11.4		WHill	7.27	36.4%		1.32	7.08%	1.33	7.12%
		13	42.7%	74.49		WHill	47.27	36.4%		8.23	6.76%	8.24	6.77%
21					**85.89**				**54.54**	**9.56**	**6.80%**	**9.57**	**6.81%**
08/17	Intr	17	42.7%	97.41		B365	61.82	36.4%		10.78	6.77%	10.77	6.76%
		70	42.6%	401.8		B365	254.55	36.4%		43.66	6.65%	43.65	6.65%
		2	42.5%	11.5		B365	7.27	36.4%		1.22	6.51%	1.23	6.55%
		100	4.24	576		B365	363.64	36.4%		60.37	6.42%	60.36	6.42%
22					**1086.71**				**687.28**	**116.03**	**6.54%**	**116.01**	**6.54%**
08/17	Intr	15	43.0%	85.5		B365	54.55	36.4%		9.96	7.11%	9.95	7.10%
23					**85.5**				**54.55**	**9.96**	**7.11%**	**9.95**	**7.10%**
08/18	Intr	15	43.0%	85.5		WHill	54.55	36.4%		9.96	7.11%	9.95	7.10%
24					**85.5**				**54.55**	**9.96**	**7.11%**	**9.95**	**7.10%**
08/19	Intr	60	42.2%	346.8		Pinn	215.83	36.0%		37.38	6.64%	37.37	6.64%
25					**346.8**				**215.83**	**37.38**	**6.64%**	**37.37**	**6.64%**
08/22	Intr	19	42.2%	109.82		Pinn	67.62	35.6%		12.57	7.09%	12.56	7.08%
26					**109.82**				**67.62**	**12.57**	**7.09%**	**12.56**	**7.08%**
08/24	Intr	2	42.4%	11.52		Pinn	7.09	35.5%		1.38	7.44%	1.39	7.47%
		21	42.3%	121.17		Pinn	74.47	35.5%		14.37	7.34%	14.36	7.34%
		513	42.2%	2965.14		Pinn	1819.15	35.5%		345.71	7.23%	345.71	7.23%
		1	42.1%	5.79		Pinn	3.55	35.5%		0.67	7.18%	0.66	7.07%
27					**3103.62**				**1904.26**	**362.13**	**7.23%**	**362.12**	**7.23%**
08/24	Intr	8	42.1%	46.32		Pinn	28.37	35.5%		5.31	7.11%	5.31	7.11%
28					**46.32**				**28.37**	**5.31**	**7.11%**	**5.31**	**7.11%**
Surebets on 1:Romney													
07/11	Intr	5	41.5%	29.25		WHill	19.05	38.1%		1.71	3.53%	1.70	3.52%
		16	41.4%	93.76		WHill	60.93	38.1%		5.25	3.39%	5.31	3.43%
		24	41.3%	140.88		WHill	91.42	38.1%		7.68	3.30%	7.70	3.31%
29					**263.89**				**171.4**	**14.64**	**3.36%**	**14.71**	**3.38%**
07/24	Intr	62	40.0%	372		WHill	236.19	38.1%		11.81	1.94%	11.81	1.94%
30					**372**				**236.19**	**11.81**	**1.94%**	**11.81**	**1.94%**
07/25	Intr	28	40.1%	167.72		WHill	106.67	38.1%		5.62	2.05%	5.61	2.04%
31					**167.72**				**106.67**	**5.62**	**2.05%**	**5.61**	**2.04%**
08/17	Intr	61	42.6%	350.14		WHill	221.82	36.4%		38.05	6.65%	38.04	6.65%
32					**350.14**				**221.82**	**38.05**	**6.65%**	**38.04**	**6.65%**
08/17	Intr	15	42.6%	86.1		Inter	54.55	36.4%		9.36	6.66%	9.35	6.65%
		60	42.5%	345		Inter	218.45	36.4%		37.29	6.62%	36.55	6.49%
33					**431.1**				**273**	**46.65**	**6.63%**	**45.90**	**6.52%**
08/26	Intr	62	43.8%	348.44		WHill	225.45	36.4%		46.10	8.03%	46.11	8.03%

34					**348.44**				**225.45**	**46.10**	**8.03%**	**46.11**	**8.03%**
08/27	Intr	112	43.9%	628.32		Lad	389.57	34.8%		102.12	10.03%	102.11	10.03%
35					**628.32**				**389.57**	**102.12**	**10.03%**	**102.11**	**10.03%**
09/09	Intr	112	42.4%	645.12		Lad	361.3	32.3%		113.61	11.29%	113.58	11.29%
36					**645.12**				**361.3**	**113.61**	**11.29%**	**113.58**	**11.29%**
10/04	Intr	51	33.3%	340.17		WHill	113.33	22.2%		56.49	12.46%	56.50	12.46%
37					**340.17**				**113.33**	**56.49**	**12.46%**	**56.50**	**12.46%**
10/23	Intr	88	41.7%	513.04		Lad	293.33	3.3%		73.62	9.13%	73.63	9.13%
		112	41.5%	655.2		Lad	373.33	3.3%		91.46	8.89%	91.47	8.89%
		52	41.3%	305.24		Lad	173.34	33.3%		41.44	8.66%	41.42	8.65%
38					**1473.48**				**840**	**206.52**	**8.93%**	**206.52**	**8.93%**
11/04	Intr	46	35.4%	297.16		Lad	106.15	23.1%		56.67	14.05%	56.69	14.06%
39					**297.16**				**106.15**	**56.67**	**14.05%**	**56.69**	**14.06%**

Notes: This table provides details on the experiment’s trades. LaySize contains the number of contracts bought on Intrade. Chance corresponds to the winning probability implied by the offered prices. LayInvest and BackInvest show the USD amount invested on Intrade or against a Bookmaker, respectively. The resulting profit and ROI are, according to the definition of a surebet, independent of the election’s outcome.

A.2 Intrade's Near Insolvency in 2013

Having closed its services to U.S. citizens in January 2013 due to legal requirements, Intrade's Board of Directors announced on its website on March 10, 2013 that:

> With sincere regret we must inform you that due to circumstances recently discovered we must immediately cease trading activity on www.intrade.com.

Later, on April 5, 2013, Intrade's director Ronald Bernstein reported:

> We have now concluded the initial stages of our investigations about the financial status of the Company, and it appears that the Company is in a cash 'shortfall' position of approximately $700,000 when comparing all cash on hand in Company and Member bank accounts with Member account balances on the Exchange system. [...] We are now very confident about the reasons which caused the current circumstance of the Company; however, for legal reasons we are not yet at liberty to document them to you. I can confirm that the Company, if it is able, intends to vigorously pursue two substantial monetary claims against two distinct parties for an aggregate amount greater than $3,500,000. Because of these circumstances, the Company has now contacted all members with account balances greater than $1,000, and proposed a 'forbearance' arrangement between these members and the Company, which if sufficient members agree, would allow the Company to remain solvent.

Finally, Intrade was able to remain solvent and handle the withdrawal requests, but the marketplace has subsequently remained inactive.

Bibliography

Arrow, K. J., Forsythe, R., Gorham, M., Hahn, R., Hanson, R., Ledyard, J. O., Levmore, S., Litan, R., Milgrom, P., Nelson, F. D., Neumann, G. R., Ottaviani, M., Schelling, T. C., Shiller, R. J., Smith, V. L., Snowberg, E., Sunstein, C. R., Tetlock, P. C., Tetlock, P. E., Varian, H. R., Wolfers, J. and Zitzewitz, E. (2008). The Promise of Prediction Markets. *Science*, **320** (5878), 877–878.

Bodie, Z., Kane, A. and Marcus, A. J. (2011). *Investments*. McGraw-Hill, 9th edn.

Borghesi, R. (2007). Price Biases in a Prediction Market: NFL Contracts on Tradesports. *Journal of Prediction Markets*, **1** (3), 233–253.

— (2011). Price Biases and Contract Design: Lessons from Tradesports. In L. Vaughan Williams (ed.), *Prediction Markets: Theory and Applications*, Routledge, pp. 96–110.

Chincarini, L. and Contreras, C. (2010). Home Bias and International Betting Markets: Can Institutional Constraints and Behavioral Biases Lead to Arbitrage Profits? *The Journal of Gambling Business and Economics*, **4** (3), 20–30.

Croxson, K. and Reade, J. J. (2011). *Exchange vs Dealers: A High-Frequency Analysis of In-Play Betting Prices*. Discussion Papers 11-19, Department of Economics, University of Birmingham.

— and — (2014). Information and Efficiency: Goal Arrival in Soccer Betting. *The Economic Journal*, **124** (575), 62–91.

De Long, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1990). Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *The Journal of Finance*, **45** (2), 379–395.

Dixon, M. J. and Pope, P. F. (2004). The Value of Statistical Forecasts in the UK

Association Football Betting Market. *International Journal of Forecasting*, **20** (4), 697–711.

Edelman, D. C. and O'Brian, N. R. (2004). Tote Arbitrage and Lock Opportunities in Racetrack Betting. *The European Journal of Finance*, **10** (5), 370–378.

Franck, E., Verbeek, E. and Nüesch, S. (2013). Inter-market Arbitrage in Betting. *Economica*, **80** (318), 300–325.

Gil, R. and Levitt, S. D. (2007). Testing the Efficiency of Markets in the 2002 World Cup. *Journal of Prediction Markets*, **1** (3), 255–270.

Griffith, R. M. (1949). Odds Adjustments by American Horse-Race Bettors. *The American Journal of Psychology*, **62** (2), 290–294.

Hausch, D. B. and Ziemba, W. T. (1990). Arbitrage Strategies for Cross-Track Betting on Major Horse Races. *The Journal of Business*, **63** (1), 61–78.

Lane, D. and Ziemba, W. T. (2004). Jai Alai Arbitrage Strategies. *The European Journal of Finance*, **10** (5), 353–369.

Levitt, S. D. (2004). Why are Gambling Markets Organised so Differently from Financial Markets? *The Economic Journal*, **114** (495), 223–246.

Oikonomidis, A. and Johnson, J. (2011). Who Can Beat the Odds? The Case of Football Betting Reviewed. In L. Vaughan Williams (ed.), *Prediction Markets: Theory and Applications*, Routledge, pp. 204–220.

Paton, D. and Vaughan Williams, L. (2005). Forecasting Outcomes in Spread Betting Markets: Can Bettors Use 'Quarbs' to Beat the Book? *Journal of Forecasting*, **24** (2), 139–154.

Rhode, P. and Strumpf, K. (2006). *Manipulating Political Stock Markets: A field Experiment and a Century of Observational Data*. Natural Field Experiments 00325, The Field Experiments Website.

Rhode, P. W. and Strumpf, K. S. (2004). Historical Presidential Betting Markets. *Journal of Economic Perspectives*, **18** (2), 127–141.

Rothschild, D. M. (2009). Forecasting Elections. *Public Opinion Quarterly*, **73** (5), 895–916.

— and Pennock, D. M. (2014). The Extent of Price Misalignment in Prediction Markets. *Algorithmic Finance*, **3** (1-2), 3–20.

— and Sethi, R. (2014). *Trading Strategies and Market Microstructure: Evidence from a Prediction Market (August 23, 2014)*. Available at SSRN: http://ssrn.com/abstract=2322420 or http://dx.doi.org/10.2139/ssrn.2322420.

Smith, M. A., Paton, D. and Vaughan Williams, L. (2006). Market Efficiency in Person-to-Person Betting. *Economica*, **73** (292), 673–689.

Snowberg, E. and Wolfers, J. (2010). Explaining the Favorite-Long Shot Bias: Is it Risk-Love or Misperceptions? *Journal of Political Economy*, **118** (4), 723–746.

—, — and Zitzewitz, E. (2012). *Prediction Markets for Economic Forecasting*. CAMA Working Papers 2012-33, Centre for Applied Macroeconomic Analysis, Crawford School of Public Policy, The Australian National University.

Spann, M. and Skiera, B. (2009). Sports Forecasting: A Comparison of the Forecast Accuracy of Prediction Markets, Betting Odds and Tipsters. *Journal of Forecasting*, **28** (1), 55–72.

Tetlock, P. (2004). *How Efficient are Information Markets? Evidence from an Online Exchange*. Working Draft, Department of Economics, Harvard University.

Vlastakis, N., Dotsis, G. and Markellos, R. N. (2009). How Efficient is the European Football Betting Market? Evidence from Arbitrage and Trading Strategies. *Journal of Forecasting*, **28** (5), 426–444.

WilliamHill (2012). Terms and Conditions (published 03/12/2008, updated 20/04/2012).

Wolfers, J. and Zitzewitz, E. (2004). Prediction Markets. *Journal of Economic Perspectives*, **18** (2), 107–126.

Woodland, L. M. and Woodland, B. M. (2003). The Reverse Favourite-Longshot Bias and Market Efficiency in Major League Baseball: An Update. *Bulletin of Economic Research*, **55** (2), 113–123.

Chapter II

Calling LABUYLA's Bluff: Empirical Evidence on a Cheating Auctioneer

Abstract. This paper analyzes the modified price reveal auction mechanism used by LABUYLA. Unusual for a pay-to-bid auction format, labuyla.ch provides sufficient public information for potential bidders to calculate the auctioned article's current hidden price for free. However, intense bidding activities indicate that most bidders are not aware of this valuable information. I compare the empirically observed behavior with the results of a model which adopts the framework of LABUYLA's auction design and includes asymmetrically informed bidders. My analyses and investigations lead to only one possible conclusion: The auctioneer was cheating by participating in the auctions as seller, bidder and buyer.

Keywords: pay-per-bid auctions; countdown clock; asymmetric information; hidden price; in-auction fraud

Jelclass: D44

1 Introduction

The LABUYLA Holding AG (registered office in Switzerland) sold selected new electronic products on labuyla.ch by means of a slightly modified price reveal auction. I provide empirical evidence that LABUYLA used thousands of fake accounts to participate in their own auctions as seller, bidder and buyer. After disclosing that I had called their bluff by revealing the source of their innovative auction design's success, LABUYLA rapidly shut down their auction platform.

In a price reveal auction, the auction house itself auctions off its own products. An auction starts at a publicly announced price, which typically equals the item's retail price. The current sales price is not publicly observable. However, it is known that the current sales price is non-increasing over time and decreases by a fixed amount for each submitted bid. The term 'bid' might be misleading, since in the context of price reveal auctions, bidding actually means that the bidder pays a fee to the auctioneer. In return, this bidder becomes privately informed about the hidden price and has an option to buy the item immediately at the revealed current price p_t. As soon as one bidder buys the item at the revealed price, the auction ends.

The hidden price has a key function in a price reveal auction. The value of the information on the current price as well as the value of the option to purchase the item are the bidders' incentives to bid and, therefore, to pay the requested fee. Since this fee exceeds the price decrease per submitted bid, the auctioneer's profit rises as the sales price decreases.

The auction design used by LABUYLA is a modification of the typical price reveal auction, such as, for example, the ones provided by dealwonders.com or dubLi.com. Detailed links and information on these platforms' availability are given in Appendix A.6. The modifications concern two crucial points in particular: Firstly, labuyla.ch publicly displays the nickname of the last bidder for all ongoing auctions. This information can be used to determine p_t without any money transaction taking place and was the stimulus for the present research. Secondly, every auction has an associated countdown clock that is independent of the number of bids submitted. When the clock expires, the auction is closed without selling the item. Note that in such a case the auctioneer retains the revenue he receives from all submitted bids. A screenshot of an auction on labuyla.ch is given in Figure 1. This figure shows the auction type (EASY; see Section 4.1 for type descriptions), the countdown clock, the last bidder's nickname, as well as the retail price, or starting price respectively, of a particular auction.

Figure 1: Screenshot of an Auction on labuyla.ch

The unique auction design used by LABUYLA has not been discussed in the literature. I set up an asymmetric model and determine the optimal strategy of a bidder who knows p_t during the whole auction period. I validate the major assumptions and the suggested optimal strategy by comparing the model's results with the empirical behavior on the website. In order to do so, I recorded the bidders' actions on the website for over three months and show how an individual becomes informed about an item's current price for free.

The behavioral analyses provide evidence that the largest share of the observed actions were performed by accounts owned by LABUYLA and, therefore, provide evidence on a cheating auctioneer.

The remainder of the paper is structured as follows. In Section 2, I discuss the related literature on pay-per-bid auctions. Section 3 provides a model of an asymmetric price reveal auction with a 'hard close' ending rule and studies the optimal strategy of a bidder who can observe p_t for free. Section 4 discloses the origin of the data and summarizes the dataset. Section 5 discusses the model's assumptions regarding the observable behavior. In Section 6, I provide evidence that LABUYLA participated in their own auctions as seller, bidder and buyer. While Section 7 concludes this paper's findings, Section 8 shares a brief anecdote about the consequences that these research results had for LABUYLA's operations.

2 Literature

The rising popularity of online auctions has encouraged various studies on fraudulent forms of behavior that bidders, sellers or the auctioneers have adopted. Several studies (Aleem and Antwi-Boasiako, 2011; Ockenfels *et al.*, 2006; Dong *et al.*, 2009; Jenamani *et al.*, 2007; Chua and Wareham, 2008) point out different types of fraud and discuss how to detect and prevent them. These investigations show that fraud prevention often focuses on e-commerce businesses that are similar in form to eBay and conclude that the auctioneer is situated in the best position to fight this type of fraudulent behavior (Snyder, 2000; Ockenfels *et al.*, 2006). If, however, an auction house itself is the cheat, then it is much more difficult to detect and reduce the fraud. One branch of this literature investigates in-auction fraud; i.e., fraudulent behavior that occurs during a running auction. Phantom betting (Wang *et al.*, 2000; Graham *et al.*, 1990; Deltas, 1999) is such an example and describes an activity where the auctioneer places a bid when only one bidder is left in the auction and thus places a flexible reserve price.

In pay-per-bid auctions, as the name suggests, bidders have to pay a fee to the auctioneer for each bid they place. The three most commonly used types are:

Penny auctions open at a ridiculously low price. However, outbidding the standing bid by only one penny is tied to a fee. Such auctions use a 'soft close' ending rule with an automatic run-time expansion for every subsequent bid. Finally, a typical penny auction ends with one lucky winner who buys the item at a very low sales price and numerous unsuccessful bidders who suffer a monetary loss. Robinson *et al.* [2013] provide an introduction to penny auctions and discuss, among other things, opportunities for fraud. Penny auctions have inspired several research papers, such as Platt *et al.* [2010] and Augenblick [2014], which theoretically and empirically analyze this non-traditional auction design. Byers *et al.* [2010] extend the literature by introducing asymmetric bidders' beliefs about model parameters.

Lowest unique bid auctions (e.g., Rapoport *et al.*, 2009) use a 'hard close' ending rule where each bidder can submit costly bids before a publicly announced deadline is reached. At the auction's end, the bidder who submitted the lowest unique bid wins the item. The sales price is typically very low and the winner makes a good bargain. Nonetheless, numerous unsuccessful bidders suffer a monetary loss. Östling *et al.* [2011] provide a theoretical and empirical analysis of the bidding behavior in a similar auction design where the lowest unique positive integer submitted wins the auction.

Price reveal auctions are descending price auctions which open at a reasonable retail price. In contrast to a Dutch auction, the standing sales price depends on the number of previously submitted bids and is not solely a function of time. To attract bids, the current price is hidden. By placing a costly bid, which is simply a euphemism for paying a fee, the bidder becomes privately informed about the current sales price. He then has, for a limited time, the option to buy the item. The final sales prices are typically not as low as in other pay-per-bid auctions. However, price reveal auctions stand out from other pay-per-bid auctions, because the uncertainty is limited to the hidden sales price. Every bidder has the opportunity to end the auction whenever he places a bid and receives the signal on the current price. The auctioneer profits from the fact that the price reduction from a submitted bid is lower than the collected fee which the bidder paid to place the bid. Gallice [2010] sets up the theoretical framework for the price reveal auction and investigates some of its properties.

All these pay-per-bid auctions share the feature that the auctioneer does not really care about the item's final sales price, because, on average, his collected fees by far exceed the given discount. In addition, these auctions seem to be particularly susceptible to cheating auctioneers who could place the winning bid themselves. These auction types operate in a gray area and their selling mechanism is legally questionable. In Germany as well as in Switzerland, penny auctions are, for example, classified as lotteries and are illegal in their present form (e.g., VGH Baden-Württemberg, 2013).

The present paper was additionally inspired by the book *Snipers, Shills, and Sharks: eBay and Human Behavior* by Steiglitz [2007] in which Steiglitz analyzes observable behavior on eBay from an auction theory perspective.

3 Model of an Asymmetric Price Reveal Auction with a 'Hard Close' Ending Rule

This model is based on the auction design used on labuyla.ch. In this section's analysis of an asymmetric price reveal auction, I assume that one bidder observes an item's current price for free and dub him the 'informed bidder'. All other bidders have to buy the information on the current price from the auctioneer by placing bids. I justify this assumption by the following two facts: (1) I show in Subsection 5.2 that any attentive observer can become informed about p_t without expense; and (2), I show in Subsection 5.3 that the intense bidding activity on the platform implies that almost certainly no bidder has used this information. Investigating the reasons why the observed bidders pay for the information on p_t rather than calculating this value for free is beyond the scope of this paper. I will treat these individuals as 'noise bidders'. In the finance literature, irrational noise traders can provide needed liquidity to the market. Bloomfield *et al.* [2009] provide an overview of the different interpretations of the term 'noise traders' and the motivation behind their behavior. The presence of noise bidders is a major prerequisite for the success of this auction design. If every bidder would be informed free of charge about p_t, then no one would have an incentive to bid and, therefore, the price would never decrease.

In this paper, I investigate the decision problem of an informed bidder who participates in an auction solely with noise traders. In order to increase comparability with Gallice's work, I use some of his notation. This section is structured as follows. In Subsection 3.1, I set up the theoretical framework. Subsection 3.2 determines the informed bidder's optimal strategy. Subsection 3.3 summarizes the model implications. Finally, Subsection 3.4 provides two model extensions, each of which addresses one assumption-driven implication of the derived informed bidder's optimal strategy.

3.1 Theoretical Framework

A single item is sold in an asymmetric price reveal auction with a hard close ending rule. The auction ends if the item is either sold, or the deadline is reached. The variable t stands for the number of seconds left until the deadline of the auction, and T stands for the initial runtime (in seconds). The current price p_t is not public information for all $t < T$. The auctioneer publicly announces four items of information: The initial price p_T, the remaining duration of the auction t, the cost c for the privately issued information on p_t, and, finally, the amount Δ that p_t is lowered for every subsequent bid.

The current price is determined by the linear function $p_t = p_T - \Delta\eta_t$, where η_t is the number of times that the hidden price was revealed at a specific point in time. To prevent a profitable strategy by which one bidder drives the current price down to zero from being exploited, $c > \Delta > 0$ must hold. By assuming that p_T is a multiple of Δ, zero is a possible current price and ensures that that p_t takes no negative values.

Bidder i assigns a valuation of X_i to the object. Each X_i is non-negative as well as independently and identically distributed according to the increasing distribution function F. Bidder i knows the realization x_i of X_i and that other bidders' valuations are independently distributed according to F.

The bidders' behavior can be characterized by four actions, namely: *wait*, in order to observe the auction process while maintaining the option to place a bid at a later time; *bid*, in order to reveal the hidden price and receive the opportunity to buy the item; *buy* the item at the current price and thereby ending the auction; and, permanently *exit* an auction.

In more detail, playing *bid* at time t has four implications, which instantly take place:

1. The bidder pays c to the auctioneer.

2. The bidder becomes privately informed about $p_t - \Delta$.

3. The bidder can either *buy* the item at the revealed price and thereby end the auction, or decline this option and, instead, either *wait* or *exit*.

4. The displayed name of the last bidder in this auction changes.

Two types of bidders who are asymmetrically informed about the current price participate in this auction. On the one hand, an informed bidder knows p_t for free during the whole auction period. On the other hand, the current price is unknown to all other bidders except for the moment t when they play *bid*.

The following three assumptions aim to reduce the length of the informed bidder's optimization problem. I discuss the model assumptions regarding the empirical behavior in Section 5.

Sequential order of actions prevents the need for a tie-breaking rule and for considering simultaneous actions. I assume that the informed bidder is able to react promptly to any price changes and that he has the ability to buy the item at the last moment before the auction closes ($t = 0$). Noise bidders arrive according to any discrete probability distribution where the probability of having two or more arrivals at the same time is zero; as given for example, by a Poisson distribution [Ross, 2010, page 316].

Deterministic behavior of the noise bidders excludes strategic decisions. An incoming noise bidder draws his private valuation x_i. He always plays *bid* and becomes informed about the current price $p_t - \Delta$ at which he can buy the item. If $p_t - \Delta \leq x_i$, he takes the action *buy*. Otherwise, he exits the market. This assumption implies that noise bidders never take the option *wait*.

Exactly one informed bidder implies, along with the noise bidders' deterministic behavior, that the informed bidder faces a decision problem instead of a game with strategic interaction.

3.2 Strategy of the Informed Bidder

The informed bidder does not gain any new information by playing *bid*. Therefore, he is able to determine in advance which of the action combinations *bid and buy*, *bid and wait*, or *bid and exit* is best. As a result of the assumption $c > \Delta$ and the deterministic behavior of the noise bidders, the informed bidder's action *bid and wait* is strictly dominated by *wait* at any point in time. The action *wait* weakly dominates *exit*, as no time costs occur in the model and playing *wait* until the auction ends is thus equal to *exit*. However, by playing *wait*, the bidder retains the option of taking the action *bid and buy* at a later stage.

For each combination of the current price p_t and the remaining time t, the optimal strategy is the answer to the question: Which of the remaining actions, either *wait* or *bid and buy*, maximizes the informed bidder's expected profit? If the informed bidder plays *bid and buy*, he receives the payoff $x_i - p_t - \bar{c}$, where $\bar{c} \equiv c - \Delta$. Playing *wait* results, at least in the short run, in one of the following three situations: (1) No noise bidder arrives; (2) A noise bidder arrives and plays *bid and buy*; or, (3) A noise bidder arrives and plays *bid and exit*.

In order to describe the informed bidder's decision problem, I define three probabilities:

$\boldsymbol{\pi_1(n,t)}$ is the probability that exactly n noise bidders arrive in the remaining t seconds of the auction.

$\boldsymbol{\pi_2(n,t)}$ is the probability that at least n noise bidders arrive in the remaining t seconds of the auction.

$\boldsymbol{\pi_3(n,p_t)}$ is the probability that none of the next n noise bidders buy the item with respect to the current price they reveal.

Note that: (1) $\pi_2(n,t) = 1 - \sum_{i=0}^{n-1} \pi_1(i,t)$; (2) $\pi_3(n,p_t) = \prod_{i=1}^{n} F(p_t - i\Delta)$, $\pi_2(n+1,t) = \pi_2(n,t) - \pi_1(n,t)$; and, (3) $\pi_3(n+1,p_t) = \pi_3(n,p_t)F(p_t - (n+1)\Delta)$.

For all combinations of p_t and t, the highest price at which the informed bidder decides to play *bid and buy* is denoted as his optimal buy-price $b^*_{p_t,t}$. The optimal strategy is directly derived from the optimal buy-price, because $b^*_{p_t,t} < p_t$ implies the action *wait*, and $b^*_{p_t,t} \geq p_t$ implies *bid and buy*. I do not use an index on the potential buy-price b, because this value always refers to the informed bidder.

Suppose that $p_t \in (\underline{p}, \bar{p})$, where $\bar{p} \equiv x_i - \bar{c}$ and $\underline{p}$ is the price that solves $(x_i - p_t - \bar{c}) = \pi_3(1, p_t)(x_i - p_t - \bar{c} + \Delta)$. In this range the current price is most interesting. If $p_t \leq \underline{p}$, then the informed bidder's optimal action is to play *bid and buy*, because the risk that the next noise bidder purchases the item outweighs the potential gain of a lower price. In contrast, if $p_t \geq \bar{p}$, then the informed bidder's optimal action is to wait for at least one further bidder, because a purchase is not profitable at the current price.

The current price decreases in discrete steps to a minimum of zero. I define $n_b\left(p_t\right) \equiv \max\left(0, \frac{p_t - b_{p_t,t}}{\Delta}\right)$, which is the minimum number of bids needed to reduce the current price such that $b_{p_t,t} \geq p_t - n_b\left(p_t\right)\Delta$ holds. Furthermore, $n_x(p_t)$ is the lowest positive integer that fulfills $\bar{p} - \left[p_t - n_x\left(p_t\right)\Delta\right] \geq 0$. As long as fewer than $n_x\left(p_t\right)$ bidders arrive, the informed bidder is better off by not buying the item. The expected profit of the informed bidder for any possible price $b_{p_t,t}$ is given by:

$$
\begin{aligned}
u(p_t, t, b_{p_t,t}) =& \pi_2(n_b, t)\left\{\pi_3(n_b, p_t)\left[\bar{p} - b_{p_t,t}\right] + \left[1 - \pi_3(n_b, p_t)\right] \cdot 0\right\} \\
&+ \sum_{i=n_x}^{n_b-1} \pi_1(i, t)\left\{\pi_3(i, p_t)\left(\bar{p} - p_t + i\Delta\right) + \left[1 - \pi_3(i, p_t)\right] \cdot 0\right\} \\
&+ \left[1 - \pi_2(n_x, t)\right] \cdot 0 \\
=& \pi_2(n_b, t)\pi_3(n_b, p_t)\left(\bar{p} - b_{p_t,t}\right) \\
&+ \sum_{i=n_x}^{n_b-1} \pi_1(i, t)\pi_3(i, p_t)\left(\bar{p} - p_t + i\Delta\right). \qquad (1)
\end{aligned}
$$

The first line represents the expected profit if at least n_b uniformed bidders arrive before the auction ends. If the first n_b of them played *bid and exit*, then the informed player could buy the item at price $b_{p_t,t}$ and gain the profit $\bar{p} - b_{p_t,t}$. The second line captures the cases where the current price does not reach $b_{p_t,t}$ until the end of the auction, but where at least n_x noise bidders arrived. If none of these noise bidders played *bid and buy*, then the informed bidder should buy the item at $t = 0$. The third line completes the possible scenarios, where the informed bidder's best choice at $t = 0$ is to not buy the item.

The informed bidder's strategy is a complete plan of action (described by $b^*_{p_t,t}$) for every possible combination of state variables p_t and t with respect to his private valuation x_i, the distribution function F, and the parameters Δ and c.

Proposition 1. *The optimal buy-price $b^*_{t>0}$ is independent of the remaining time t and the current price p_t.*

The proof is given in Appendix A.1. The situation changes at $t = 0$ when the informed bidder has a final chance to purchase the item. He should then buy the item if $\bar{p} - p_0 \geq 0$. If $b^*_{t>0} \neq b^*_{t=0}$, then the optimal buy-price will immediately jump at the last possible moment before the auction ends.

3.3 Discussion of the Model's Result

The informed bidder's willingness to pay jumps at the last moment of the auction which might seem strange at first sight. The model implies that an informed bidder should calculate his optimal buy-price $b^*_{t>0}$, which does not change until the countdown clock reaches zero. While $t > 0$, he purchases the item if $p_t \leq b^*_{t>0}$. At the last moment of the auction, when the clock reaches zero, he buys the item if the utility of the purchase is non-negative, or $p_0 \leq \bar{p}$, respectively. This result is driven by the assumptions about the deterministic behavior of the noise bidders and the infinitely fast reaction time of the informed bidder.

The jump at the last moment, or the late buy, has nothing to do with a sniping strategy (Roth and Ockenfels, 2002; Ockenfels and Roth, 2006; Ariely *et al.*, 2005; Hossain, 2008; Chakraborty and Kosmopoulou, 2004). Sniping is a strategy used to prevent other bidders from learning and adjusting their valuation of the product. However, in my model no strategic interaction between bidders exists, and bidders know their precise valuation of the item. In fact, the reason for the jump in the optimal buy-price is the option to buy the item at a later point in time, whenever no further noise bidders arrive. As long as this option exists and the arrival of an additional uniformed bidder is desirable for the informed bidder, his best action is *wait*. The option expires at the last moment of the auction, and this causes the jump in the optimal buy-price.

3.4 Model Extensions

The first extension aims to avoid the case where the informed bidder's valuation is so high that he buys the item right at the start of the auction, such that $b^*_{t>0} = p_T$. I introduce an outside option to buy the same item and assume that $2\Delta > c$ holds. The item is thereby unlimitedly available at the retail price p_T in a store and can be purchased without any additional costs. These two additional assumptions reflect the auction design used by LABUYLA.

In a second extension, the informed bidder's submission of the bid and purchase order will probably not be successful. This uncertainty prevents the counter-intuitive jump in the optimal buy-price at the last possible moment before the auction ends.

3.4.1 Outside Option to Purchase the Item at a Store

The behavior of the noise bidders is basically the same. The only difference is that if they arrive and the auction is already closed, then they buy the item at the store if $x_i > p_T$. Therefore, the informed bidder's strategy is not influenced by the change in the noise bidder's behavior. However, the informed bidder takes the introduced outside option into account if $x_i > p_T$, and either $\eta_0 = 0$ or an uniformed bidder were to have bought the item.

Proposition 2. *If $t > 0$ and $x_i > p_T$, then the additional option of buying the item at a fixed price increases the informed bidder's expected profit.*

The proof is given in Appendix A.2. If no noise bidder arrives, then the informed bidder can purchase the item at the online store at price p_T instead of buying the item at the auction at price $p_T + \bar{c}$. In addition, he can purchase the item at the online store if a noise bidder bought the item in the auction. This option leads to a lower optimal buy-price $b^*_{t>0}$. If $t = 0$, the informed bidder buys the product in the auction if the item has not been sold and at least one noise bidder has arrived; otherwise, he buys the product at the online store. This result is ensured by the assumption that $2\Delta > c$. If $x_i \leq p_T$, then the option is useless for the informed bidder, and the basic model can be used.

3.4.2 Uncertain Bid Submission

The bid of an informed bidder is still certain to be submitted if $t > \bar{\omega}$. However, suppose that if $t \in [0, \bar{\omega}]$, the bid and purchase order of an informed bidder is submitted with probability $\frac{t}{\bar{\omega}}$. I further assume that noise bidders arrive according to a Poisson process with mean λ per second.

This newly introduced uncertainty means that the action *wait* becomes costly during the last $\bar{\omega}$ seconds of the auction. Thus, if the informed bidder's best option is *wait*, how long should he wait for a noise bidder to arrive? The action *bid and buy* is compared to the wait option with the highest expected profit. I define $\omega^*_{p_t}$ as the optimal buy-time if no additional noise bidder arrives. Suppose a noise bidder arrives $z \in [0, \bar{\omega}]$ seconds before the auction's deadline and plays *bid and exit*. In this case, the expected profit of an informed bidder playing his best strategy is Ω^*_{z,p_t}, which must be at least as high as the expected profit from submitting a bid right away, or $\frac{z}{\bar{\omega}}(\bar{p} - p_t + \Delta)$, respectively.

The first term in Equation 2 represents the profit if no further noise bidder arrives and the informed bidder purchases the item ω_{p_t} seconds before the auction ends. The second term represents the expected profit if a noise bidder arrives, conditional on the point in time when he arrives.

$$\Omega_{t,p_t} = e^{-\lambda(t-\omega_{p_t})}\frac{\omega_{p_t}}{\bar{\omega}}(\bar{p} - p_t) + \int_{\omega_{p_t}}^{t} \lambda e^{-\lambda(t-z)}\pi_3(1, p_t)\Omega^*_{z,p_t-\Delta}dz. \tag{2}$$

In order to derive $\omega^*_{p_t}$, I take the following three steps: Firstly, I consider only current prices which fulfill $p_t \in (\underline{p}, \underline{p} + \Delta]$ and $t \in [0, \bar{\omega}]$. I show that in these cases $\omega^*_{p_t}$ is independent of t. Secondly, I prove that if $t \in [0, \bar{\omega}]$, then $\omega^*_{p_t}$ is generally independent of t for all $p_t \in (\underline{p}, \bar{p})$. Therefore, $\omega^*_{p_t}$ does not depend on the possibility that the informed bidder might prefer to wait for more than one additional noise bidder. Finally, I show that $\omega^*_{p_t}$ as well as the implied strategy is valid even if $t > \bar{\omega}$.

Proposition 3. *If $p_t \in (\underline{\mathrm{p}}, \underline{\mathrm{p}} + \Delta]$ and $t \in [0, \bar{\omega}]$, then $\omega^*_{p_t}$ is independent of t and strictly positive.*

The proof is given in Appendix A.3. Whenever $\omega^*_{p_t} < t$, the expected profit of the action *wait* exceeds the profit of playing *bid and buy*, and vice versa. In other words, the value of the option *wait* is positive if $\omega^*_{p_t} < t$ and equals zero if $\omega^*_{p_t} \geq t$.

Proposition 4. *If* $t \in [0, \bar{\omega}]$, *then* $\omega^*_{p_t}$ *is generally independent of* t *for all* $p_t \in (\underline{p} + \Delta, \bar{p})$.

Proposition 5. *Proposition 4 is also true for* $t > \bar{\omega}$.

The proofs are given in Appendix A.4 and A.5, respectively. Because waiting is costly, the informed bidder's willingness to pay increases as the deadline approaches. At a current price less than or equal to $\underline{p}$, the informed bidder's best action is to play *bid and buy*. If the current price exceeds $\underline{p}$, then he should wait for at least one further bidder, as long as $t > \omega^*_{p_t}$. Note that $\omega^*_{p_t} \in (0, \bar{\omega})$ and $\omega^*_{p_t}$ decreases as p_t increases. Uncertainty about the successful transaction of the bid submission eliminates the counter-intuitive jump of the informed bidder's willingness to pay close to the auction's end.

4 Empirical Data

I have collected the empirical data for two reasons: Firstly, I wanted to use the dataset to check whether the bidders' behavior appears to be non-optimized from an external perspective and thereby supports the model's assumption on the deterministic behavior of the uninformed bidders; Secondly, I wanted to show how the current price can be calculated by only using publicly available information. As this precluded the option of asking the owner about the data, I recorded all the auction activities that occurred on labuyla.ch myself.

4.1 Data Collection

The labuyla.ch start page displays the 9 auctions with the currently lowest t, while auctions with a higher current t are listed on subsequent pages. As shown in Figure 1, the product name, a picture of the item, the auction type (c and Δ), the nickname of the last bidder, and the value p_T as well as t are easily visible for each auction. Whenever a bidder reveals the hidden price, all other potential bidders can observe this action because the displayed name of the last bidder changes, accompanied by an eye-catching red flash. Precisely this succession of the displayed bidder's nicknames enabled me to collect a comprehensive dataset of the bidding activities on labuyla.ch.

A closer look at the transmitted source code of the website shows a structured HTML code. Every auction carries a unique identification number, which is specified in a specific section of the source code. Moreover, all the public information recorded in the dataset is freely available and explicitly tagged. I collected and stored the data by using an algorithm programmed in JAVA. The algorithm assigned the stored values to the website's source code by processing the data in an infinite loop. If the identification number was not found in my database, all the information of the auction was saved. On the contrary, if the identification number had been previously recorded in the database, the current nickname of the last bidder was then compared with the nickname of the last bidder in the dataset. If they were different, the current nickname of the last bidder as well as t were added to the dataset.

At the end of March 2010, I noticed a change in the bidding behavior. Before this date, nearly all bids took place in the 9 auctions with the currently lowest t. Subsequent to this date, bidders revealed the hidden price in an earlier stage of the auction (higher t). My first algorithm captured only the activities in the last 9 auctions, which are displayed on the website's start page. This procedure was sufficient to record nearly all the bids performed on the platform in accordance with the old behavior. However, I had to respond to the new betting behavior by adjusting the algorithm. The adjusted algorithm also checked and recoded bids in auctions listed on the second to tenth page. The algorithm requires less than 0.2 seconds to call up the source code of a single URL, compare the information with the database, and finally save any new information (1 cycle). While the first algorithm called up the start page in every cycle, the adjusted algorithm called up the start page either every 0.2 or 0.4 seconds. On the other hand, a cycle evaluated the current information on the second page every 0.8 seconds, and the third to tenth pages were called up every 32 seconds. Whenever more than one bid on a specific auction occurred between two page requests, no more than one bid was recorded.

In addition, labuyla.ch presents all the auctions that ended with a purchase in their bestseller list. The bestseller list includes, among other things, the nickname of the buyer as well as the buy-price. I recorded and matched this list with the observed bidding behavior in the database in order to verify the algorithm's observation accuracy.

4.2 Dataset Summary

In the period from the 12th of March 2010, 18:16:03, to the 14th of June 2010, 02:48:02, I observed and recorded the bidding behavior on labuyla.ch. During this observation period, I logged 523,837 actions on the platform. These actions were either of type S (start of a new auction), BW (*bid and wait*), or BB (*bid and buy*). An auction was labeled as incomplete if at least one of the following criteria was true:

- At least one bid was already made (last bidder's name was not 'None') when the algorithm detected the auction.

- The auction took place in an interval, where the observation was interrupted.

- The buyer specified in the bestseller list did not match the last bidder recorded by the algorithm.

Incomplete auctions were eliminated from the dataset. LABUYLA runs three types of auctions: EASY ($c = 0.95$; $\Delta = 0.50$), BIG ($c = 9.50$; $\Delta = 8.00$) and CRAZY ($c = 95.00$; $\Delta = 90.00$). In the following analyses, except for Subsection 6.2, only auctions of type EASY are analyzed, which includes 14,940 of the 19,739 auctions recorded in total. Auctions of type EASY accounted for more than 95% of all purchased items and for more than 99% of all actions. This final dataset, which only contains complete auctions of type EASY, is termed τ.

4.3 The Algorithm's Observation Accuracy

Even though an auction is categorized as complete, the applied algorithm cannot guarantee – either because several bids occurred between two page requests or because a single bidder revealed the hidden price several times in a row – that all bids were recorded. The buy-price, recorded in the bestseller list, provides information on the true number of bids in this auction. Therefore, by comparing this value to the number of nickname changes detected by the algorithm, I determine the algorithm's observation accuracy. The basis for this analysis are all the auctions in the dataset τ in which the auctioned articles were sold.

Figure 2: The Algorithm's Observation Accuracy

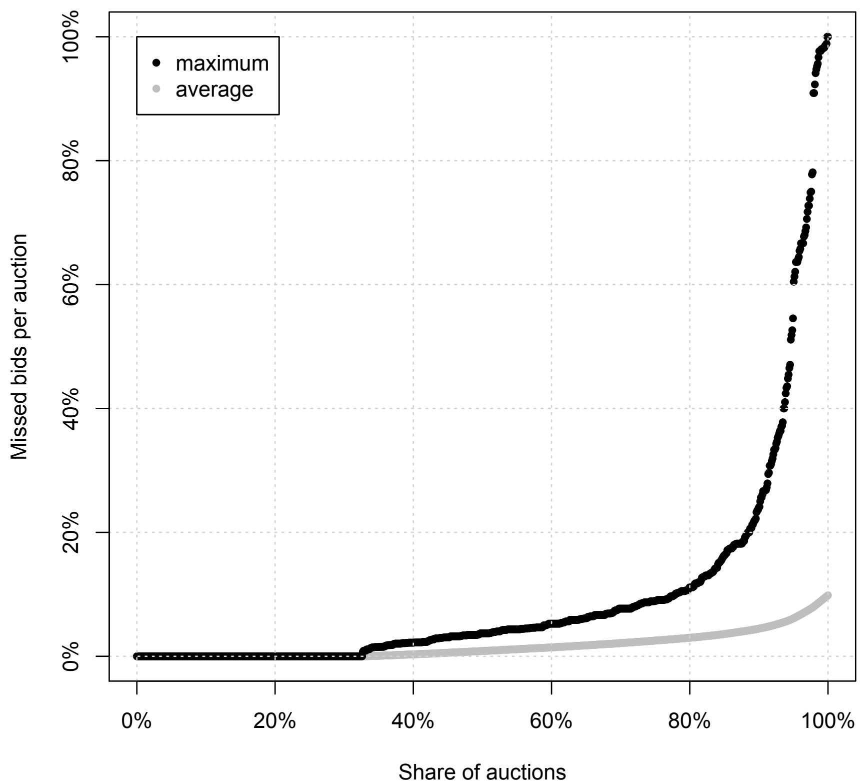


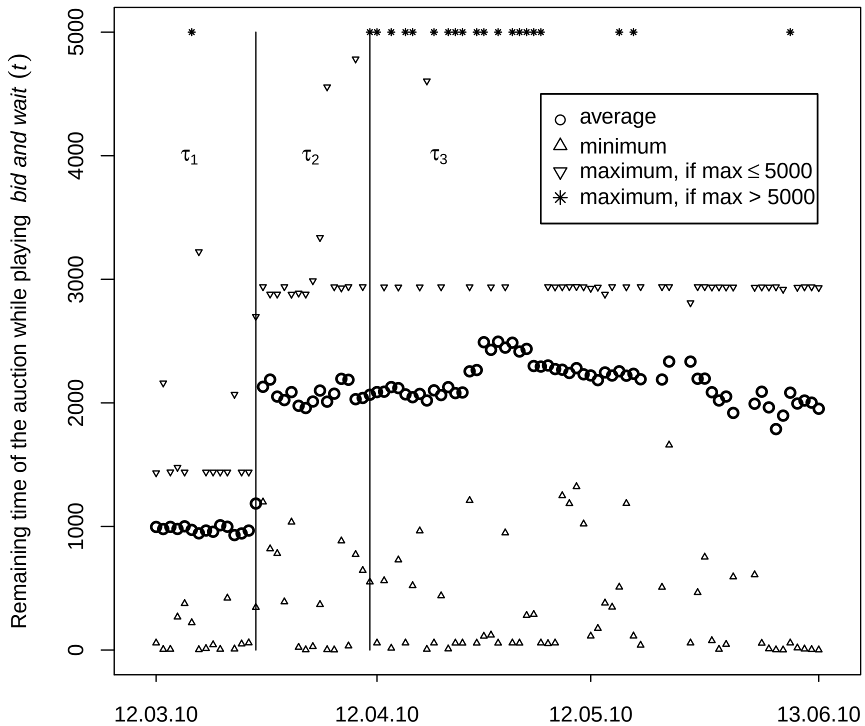
Figure 2 presents the share of observed bids in relation to the true number of bids. This figure shows that in 33% of these auctions all bids were observed, and in 90% of these auctions at least 75% of all bids were observed. On average, the algorithm missed less than 10% of all bids. Therefore, most of the submitted bids were recorded.

4.4 Classification of Subsets

As mentioned before, I had to adjust the algorithm with regard to a change in the bidding behavior. To illustrate this change, Figure 3 summarizes the daily bidding behavior.

For every day (04:00 a.m. to 03:59 a.m.) this figure shows the minimal, the maximal and the average t of the recorded actions BW. The doubling of the average, and maximal t from one day to another can be seen. After a closer look at the daily data from the end of March, I ascribe this change in behavior to the 26th of March 2010, 17:29:25. During

Figure 3: Change in Bidding Behavior and Subset Classification



this day the recording algorithm worked properly and no problems requiring mention were observed. My adjustment of the algorithm is the reason for the second partition on the 11th of April 2010, 20:10:20.

According to these two insections, I split the dataset τ into the three subsets τ_1 (*old behavior* and *first algorithm*), τ_2 (*new behavior* and *first algorithm*), and τ_3 (*new behavior* and *adjusted algorithm*). Table 1 shows the number of observed actions in each subset.

5 Discussion of the Model's Assumptions Regarding the Observable Behavior

I use the empirical dataset to discuss the model's major assumptions in Section 3.1 that are essential for evaluating the theoretical findings. In addition, I use the observed behavior to improve the strategy of an informed bidder's interaction on labuyla.ch and discuss alternative explanations of the observed bidding behavior.

Table 1: Number of Actions in Each Subset

Dataset	S	BW	BB	$\sum$
τ_1	2,298	88,496	92	90,886
τ_2	1,920	86,369	118	88,407
τ_3	10,722	293,627	737	305,086
τ	14,940	468,492	947	484,379

Notes: This table contains the number of started auctions (S), the number of bids placed without buying the item (BW), and the number of purchases (BB). The dataset τ is subdivided into τ_1, τ_2 and τ_3 with respect to the applied recording algorithm and the bidders' behavior.

The bidders' private valuations of an auctioned product are not observable. For this reason, I define the term profit in this section as follows: Profit equals the discount on purchases minus the expenditures on placing bids, and thus reflects the bidder's monetary gain from participating in auctions.

5.1 Becoming an Informed Bidder

The presented model requires the possibility of knowing the current price without any payment to the auctioneer being necessary. Due to the fact that the algorithm records the bidding histories of all currently running actions, the information indicating how many bidders have already revealed the price (η_t) is available for ongoing auctions. Adding this information to the publicly known values for Δ and p_T reveals the hidden prices on labuyla.ch.

5.2 No Other Informed Bidder

The model is designed for only one informed bidder. Thus, no other informed bidder should be present in a typical auction. To identify the bidder's type, I identify two characteristics that distinguish an 'informed' bidder from a 'noise' bidder. I classify a bidder as potentially informed, if the following two conditions are fulfilled:

1. Profit > 0: A rational and informed bidder profits from his participation in the auctions. If his expected savings from a lower price do not exceed his expenses for revealing the hidden price, he could increase his utility by buying all products at

p_T at the online store. On the contrary, the noise bidders follow the given strategy even if they make a negative profit.

2. Ratio *BW:BB* ≤ 9: In the model, an informed bidder always plays *bid and buy* or *wait* and thus this ratio should be zero. Nevertheless, I choose 9 as the critical value in order to account for the fact that on the website a bidder has to pay for at least 10 bids in advance. An informed bidder who wants to buy only one item would therefore lower the price nine times prior to his purchase.

The dataset τ contains 2,883 different bidders or nicknames, respectively. Of this total, 2,075 bidders played solely the action *bid and wait*. Because this behavior is exactly the opposite action that an informed bidder would take, these bidders are classified as noise bidders. The remaining 808 bidders bought at least one item and are classified according to their gained profit and their actions' ratio of *bid and wait* to *bid and buy*. Table 2 shows that 10 bidders feature both characteristics of an informed bidder. Another 9 bidders at least satisfy the first condition, the profit motive.

Table 2: Identifying Potentially Informed Bidders

	BW:BB ≤ 9	*BW:BB* > 9	Σ
Profit > 0	10	9	19
Profit ≤ 0	0	789	789
Σ	10	798	808

Notes: The 808 bidders who bought at least one item are grouped according to their total gains from the auction and their *BW:BB* ratio.

The 10 potentially informed bidders in τ bought a total of 12 items. Even if they are actually informed bidders, in comparison to the 14,940 recorded auctions, the probability that one of them buys a particular item is negligible. This insight supports the assumption that no other informed bidder is involved in the auctions. The result would not change by assigning the type only according to the bidder's profit, because the 19 bidders with a positive profit bought a total of only 23 items.

5.3 Deterministic Behavior of the Noise Bidders

It is certainly tenable that even a noise bidder should be able to recognize that at least one bidder has revealed the hidden price by simply checking whether the field which displays the name of the last bidder is empty. If this supposition is true, then even he should be able to conclude that the hidden price equals the commonly known price p_T. In such a model, the noise bidder would not make the initial bid, and the informed bidder might now have an incentive to place the initial bid and ensure that an arriving noise trader will play *bid*. However, such a rationale is not in line with the observed bidding behavior.

The underlying assumption that noise bidders observe the hidden price at most once per auction and play *bid* without conditioning their actions on the publicly available auction details or on their private valuations is a simplification in order to shorten the calculations without losing crucial insight. To validate this simplification, I eliminate the ten bidders, whom I classified as potentially informed, from τ for the analyzes in this subsection.

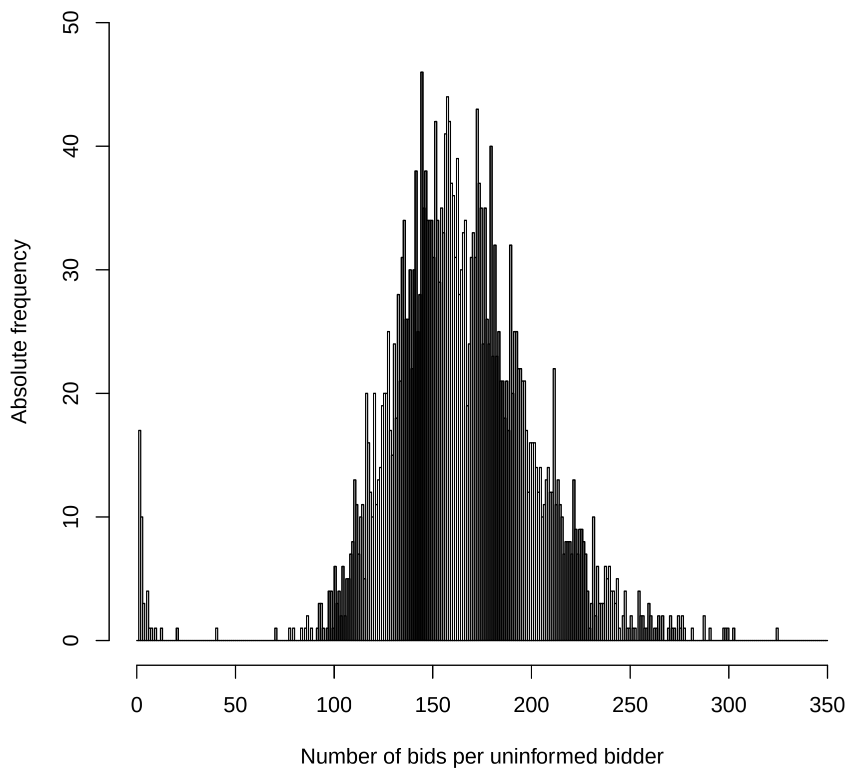
5.3.1 Number of Bids per Noise Bidder Over All Auctions

The histogram in Figure 4 shows the number of bids per noise bidder over all auctions. Whereas 37 bidders made 1 to 10 bids, and 10 bidders made 11 to 90 bids, 2,826 of the 2,883 bidders, or 98% of all noise bidders, decided to place over 90 bids, resulting in a negative profit. Aggregated over all noise bidders, the current price was revealed 469,342 times at a price of CHF 0.95. Noise bidders bought the item at the revealed price only 935 times. These facts provide no information about the decision rule or strategy used by noise bidders. However, arguable implications of the assumed noise bidder's strategy, such as negative profits or numerous noise bidders who simply lower the hidden price, are in line with this observation.

5.3.2 Number of Bids per Noise Bidder per Auction

Purchase the product or exit the market: these are the only actions allowed in the model for a noise bidder after he has revealed the hidden price. If a bidder reveals the hidden price more than once, he either bids strategically by taking potential externalities into account or does not have the sagacity to recognize the information behind the changing

Figure 4: The Number of Bids per Noise Bidder Over All Auctions



nickname of the last bidder. My assumptions prohibit multiple bids per noise bidder, as I neither believe that a noise bidder, whose behavior results in a monetary loss, bids strategically, nor that he lacks shrewdness.

Analyzing the dataset reveals that in 229,844 cases (73%) a noise bidder bid exactly once per auction. In 54,258 cases (17%) a noise bidder bid exactly twice per auction, and in 16,797 cases (5%) a noise bidder revealed the price three times per auction. This result supports the model's framework.

5.3.3 Timing of the Noise Bidders' Actions *Bid and Wait*

In the model, I do not specify the discrete probability distribution according to which noise bidders arrive. However, I consider an auction to appear more attractive when t is low, because the nine auctions with the lowest t are presented on the website's start page. As a result, they receive more attention and are likely to have more potential bidders. Along with the expectation that the hidden price of auctions with a low t has been revealed more often, I expect to observe an accumulation of bids in the final minutes of an auction. The remaining time t after a bid has been placed is represented in Figure 5 and Figure 6. Only bids which took place at a remaining auction time of less than 3,000 seconds are considered.

Figure 5: Timing of the Played Actions *Bid and Wait* in Each Subset

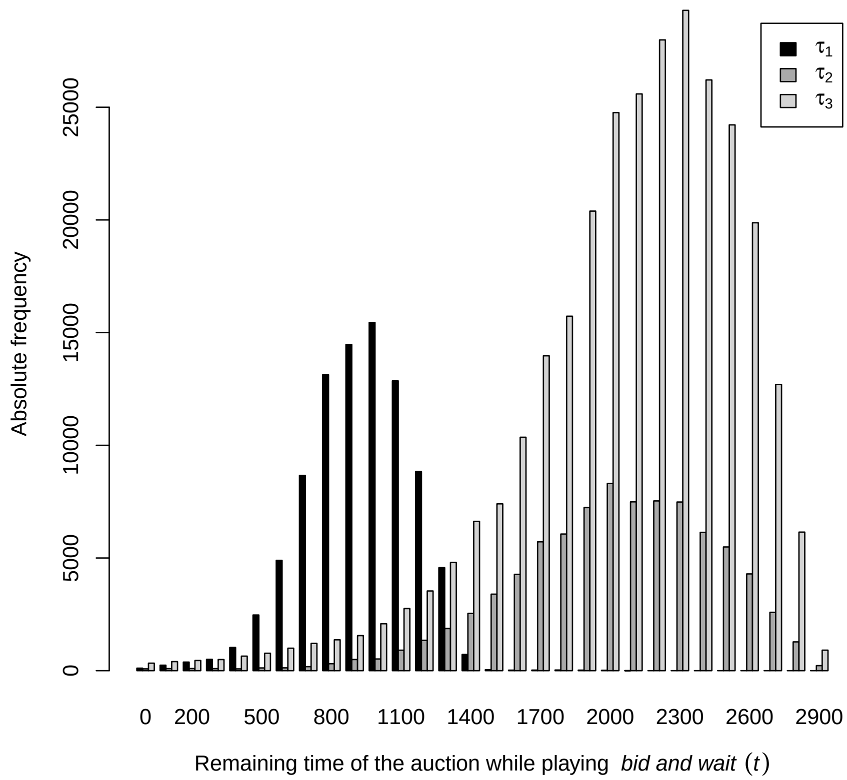
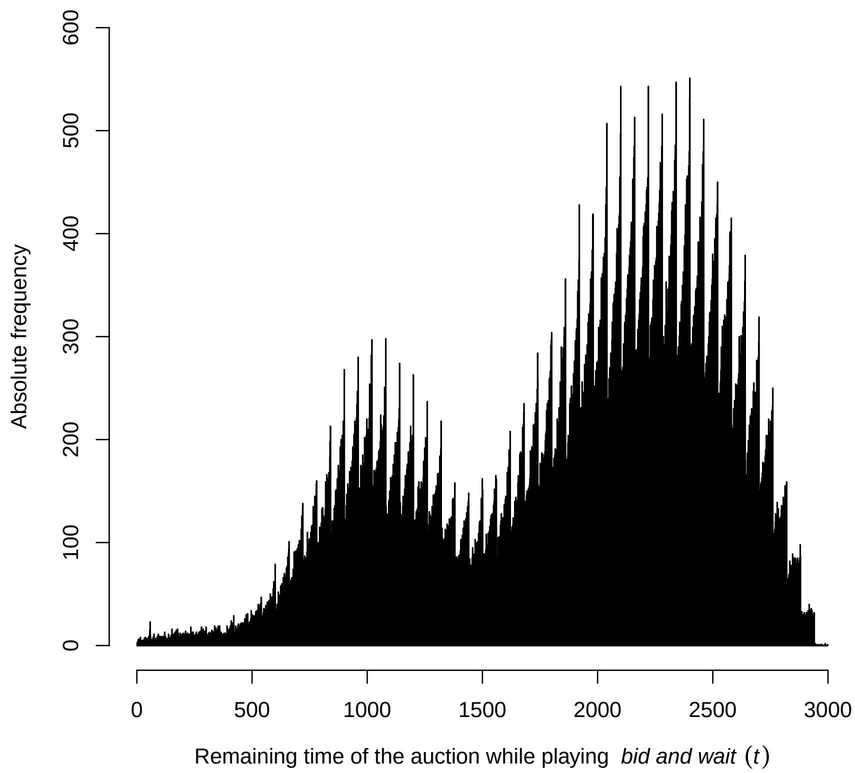


Figure 5 distinguishes bids according to the periods τ_1, τ_2 and τ_3 in which they occurred. As in Figure 3, the sudden increase in average t is apparent. Most of the bids occur around the average value, and no increase in bids appears close to the end of the auction. By focusing on the exact value of t rather than on the change in behavior, Figure 6 shows peaks at constant intervals. A closer look reveals that each interval is exactly 60 seconds, and the peaks appear in the moments where t is a multiple of 60. This mechanical pattern can be explained neither by the model nor by any adequate justification that I can identify.

The data could support the model if many items were sold when t is high, enabling

Figure 6: Number of Played Actions *Bid and Wait* for Each Value of $t \leq 3,000$

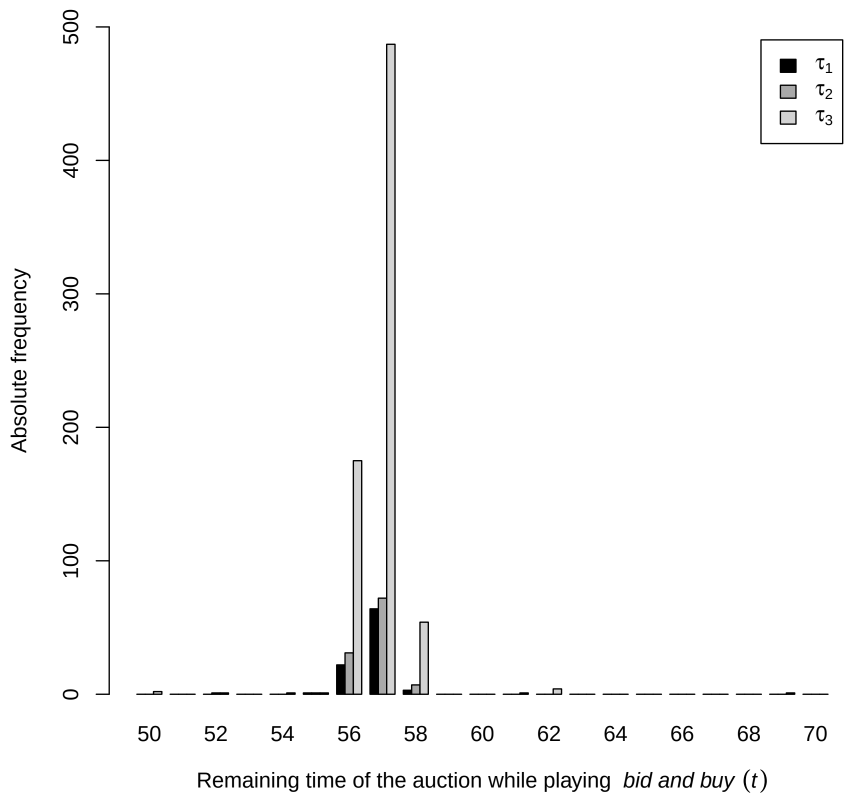


the lower count of bids at a low t to be explained by the fact that fewer auctions reach such a low t. To check this argument, Figure 7 shows the time t at which items were sold according to τ_1, τ_2 and τ_3. This figure misses 6 items that were sold when $t < 50$ and 12 items that were sold when $t > 70$.

Purchases are highly concentrated at $t = 57$ and at the adjoining seconds in all three periods. Out of a total of 947 items, 915 items were sold when the countdown clock displayed 56, 57 or 58 remaining seconds. The proposition that hundreds of noise bidders independently arrive and buy at $t \in \{56, 57, 58\}$ is unrealistic in my opinion.

To account for the conjecture that this observation is the result of a single bidder's action, I check the number of items that were bought by a single buyer, on the basis of their nickname identity. From a total of 785 buyers, 670 buyers bought one item, 100 buyers bought two items and 15 buyers bought three items. No single nickname is responsible for purchasing more than three items. I thus reject the conjecture that this

Figure 7: Timing of the Played Actions *Bid and Buy*



phenomena is the result of one single bidder's strategy.

6 Evidence on a Cheating Auctioneer

The independence of the noise bidders from the environment and their non-optimized strategy are strong assumptions of the model. The dataset reveals different patterns in the aggregated bidding behavior which cannot be explained so far. In contrast to my prediction, the behavior does not appear to be random from an outside perspective. A repeated structure and focus points are identifiable in the behavior evident in the dataset. As these data result from empirical observations, these patterns are hard to arrange subject to the independence assumption. Instead of adjusting the model to these findings, I check the data for two possibilities where the independence assumption might be violated.

Until now, I have assumed that every bidder has only one identity on the auction platform. From now on I will consider the possibility that a bidder uses multiple nicknames on labuyla.ch.

6.1 One External Bidder with Multiple Identities

In a theoretical paper, Yokoo *et al.* [2004] analyze the effects of false-name bids on combinatorial auction protocols. They state that identifying bidders who submit false bids by using multiple e-mail addresses are difficult to detect.

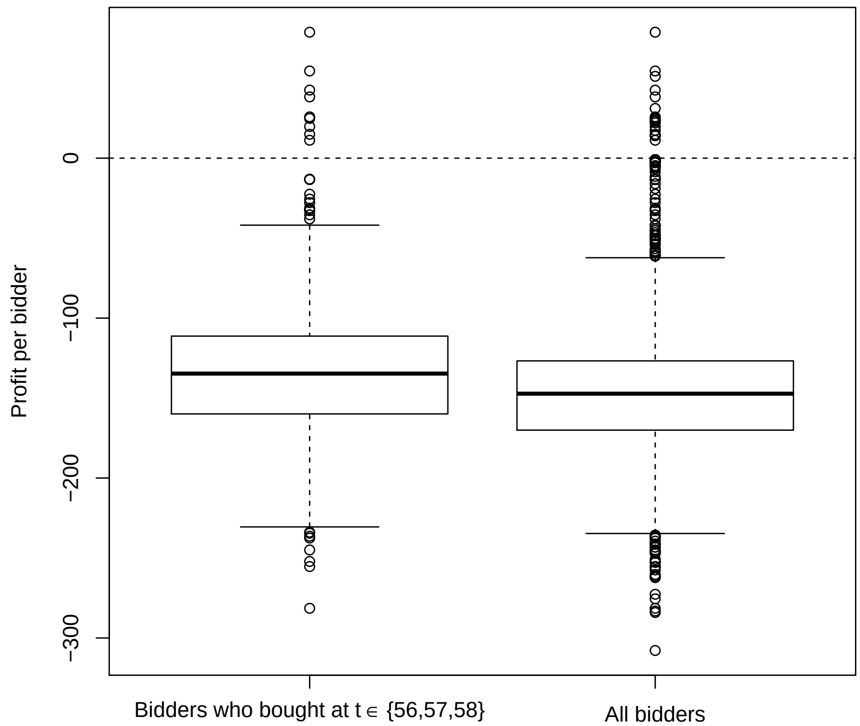
I suppose that one bidder registered more than one nickname on the auction platform, which allows him to take part in the auctions under different nicknames. Such behavior might be a reasonable strategy if this bidder gains an advantage from hiding his identity and his strategy from other bidders or from the auctioneer. If his strategy, for example, is *buy the item 57 seconds before the auction ends if the price at that moment is lower than the private valuation*, this strategy could explain the observed accumulation. Figure 8 summarizes the accounts' profit, which is calculated as the received discount on purchases minus the expenditures for bidding. The boxplot on the left-hand side contains only bidders who bought at least one product at $t \in \{56, 57, 58\}$. The boxplot on the right-hand side contains all bidders.

The setting in which one bidder exists who uses hundreds of nicknames and is responsible for the purchases, instead of hundreds of noise bidders who independently arrive and buy at the same second, is capable of explaining the observed coordinated behavior. However, two reasons cause me to reject this explanation:

1. Multiple nicknames imply different personal data and shipping addresses. Otherwise, the strategy only misleads other bidders, but the auctioneer has to tolerate the strategy.

2. The predominate negative profit, shown in Figure 8, and the large count of the actions *bid and wait* of the nicknames who purchased items at $t \in \{56, 57, 58\}$ does not present a financially rewarding strategy.

Figure 8: Profit per Bidder from Participating in the Auctions



6.2 The Auctioneer Directly Influences the Auctions

Another explanation for the observed coordinated behavior which does not face the two impediments of an external bidder with multiple identities, is an internal bidder (the auctioneer himself) with multiple identities. Three possible incentives exist which might induce the auctioneer to participate in the auction while his own nicknames make negative profits; these incentives are:

1. Playing *bid and wait* can be profitable if the auctioneer's bid attracts an observer to reveal the price. Each of his *bid and wait* actions costs the auctioneer Δ if the item is sold to an external bidder, or is otherwise free. In contrast, from each bid that is attracted by his own action, the auctioneer earns $\bar{c}$ if the item is sold to an external bidder, or c otherwise.

2. Playing *bid and buy* can also be profitable for the auctioneer. Even though he

immediately ends the auction and forgoes the possibility of earning money from noise bidders who subsequently arrive at the auction, such an action hedges the auctioneer against the risk of selling the product to a bidder at the reduced hidden price. This protection guarantees that the auctioneer gains a profit of c for all bids made by external bidders, instead of $\bar{c}$ per bid if an external bidder buys the item.

3. Another possibility why such a strategy might be profitable can be seen from an indirect perspective. As owner of the website and developer of this specific auction design, the auctioneer's intention is that the platform gives the appearance of providing an efficient service to the external observer. This appearance of legitimacy serves to attract more potential bidders and also assists the auctioneer in fulfilling strategic business objectives.

Profit in this context is the auctioneer's gains from selling the items through the auction platform instead of selling them through the online shop. I calculate his profit by subtracting the discount given for the sold items from the inflated payments gained by revealing the hidden price. The auctioneer's profit, shown in Table 3, is based on all complete auctions of all three types EASY, BIG, and CRAZY.

Table 3: Auctioneer's Profit from the Price Reveal Auctions

Auction Type	Gain (CHF)	Discount (CHF)	Profit (CHF)
CRAZY	1,140.00	720.00	420.00
BIG	17,708.00	2,787.40	14,920.60
EASY	445,967.00	26,665.40	419,301.60
Σ	464,815.00	30,172.80	434,642.20

Notes: This table shows how much money the auctioneer earned from selling via the auction rather than selling through the online store; assuming that he does not own any nicknames himself. Gain: Sum of all monetary transactions from bidders in the auctioneer as a result of the actions BW and BB. Discount: Sum of all differences $p_T - p_t$ if an item was purchased at the price p_t.

According to the dataset, the auction platform realized a profit of over CHF 430,000 during the observation period (about 3 months). The calculation of the profit is biased if the auctioneer owns nicknames. In the worst case for the auctioneer, he owned all the nicknames that yield a negative profit. In this case, he would have made a loss of CHF 777.30. The counterpart would be if he owned all the nicknames that made a positive profit, resulting in a profit of CHF 438,750.50.

The auctioneer might act even more fraudulently by playing *bid and wait* without lowering the hidden price. However, I rule out such a behavior, as the data, used for Figure 2, do not contain any case where the calculated buy-price is lower than the buy-price recorded in the bestseller list.

Even though Aleem and Antwi-Boasiako [2011] discuss many different types of fraud, they do not cover this type of behavior. LABUYLA's auction design allows the auctioneer to act even more fraudulently than by placing phantom bids. While phantom bids would lower the current price, LABUYLA can close an auction at any point in time by playing *bid and buy*.

6.3 Strategy According to the Implications of the Dataset

To outline a profit-promissory strategy where no optimized reservation price must be calculated, I use the following facts from τ (auctions with no bidders are ignored in calculating the following values):

1. The probability that an auction ends at $t > 60$ is below 2%.
2. If at least one bidder played *bid and wait*, the price p_{70} is on average 8.6% (standard deviation: 0.06) below the retail price.
3. During the period $t \leq 70$ on average 0.03 bidders arrive and play *bid and wait*.

An informed bidder should wait until $t = 70$ and then decide whether he wants to buy the item at the current price or not to bid at all. I choose the threshold at $t = 70$, as this value gives the bidder 10 seconds time to handle all the information and play his preferred action before the countdown clock reaches 60.

This strategy serves as decision guidance and not as a decision rule. Which price the bidder should respond to and which action he should take are not definitively determined.

Note that this strategy guidance is based on the observed environment and decreases in utility if more than one bidder follows the suggestions. Noise bidders who reveal the price and do not buy the item are essential for a low, and thus attractive, current price to be reached.

7 Conclusion

In this paper, I set up a model with asymmetric informed bidders who participate in the modified price reveal auction mechanism used on labuyla.ch and test the theoretical findings with regard to the empirical behavior. The record of the empirical behavior on labuyla.ch is an important part of this paper. I demonstrate the possibility of gaining information on the hidden price without any money transaction taking place. According to the dataset, of the 2,883 bidders recorded in total, a maximum of 10 are potentially informed. This strengthens the validation of the model, which assumes that no other informed bidder exists in a particular auction. Contrary to my expectation, the noise bidders' behavior is even more static and mechanical than I assumed. Strictly speaking, I cannot think of any framework, containing autonomous, real people, that is compatible with the observed behavior. Three incompatible facts sustain this argument:

1. Almost all bidders who purchased an item did so independently between 56 and 58 seconds before the countdown clock was due to expire.

2. All bidders changed their strategy from the *old behavior* to the *new behavior* within a few minutes.

3. Bidders were attracted by auctions when the current remaining countdown time was a multiple of 60.

The analysis of empirical behavior, recorded in the dataset, provides only one explanation that I could not reject: Almost all of the nicknames used are not owned by actual market participants; instead, they are owned by the auctioneer.

Even though the real motivation for such behavior can only be guessed, this explanation alone is in line with the dataset and the observed patterns. This cheating behavior of the auctioneer, who sells all the products, is a new type of fraud which allows the auctioneer to fraudulently increase his profits in nearly every circumstance.

A cheating auctioneer who operates in an anonymous, lottery-like, fast-paced and legally questionable online environment is definitely not inconceivable. However, this is the first paper to identify and provide empirical evidence for this type of cheating auctioneer and complements previously detected types of in-auction fraud.

Further research on comparable auction or selling mechanisms might, therefore, take fraudulent auctioneer behavior into consideration. In addition, this paper shows that a new auction design which is prematurely released might provide valuable information. Valuable either from research perspective, or to gain profit through an information advantage. Even though labuyla.ch provided a potentially unique modification of a price reveal auction, I am sure that this is not the only platform which suffers such misconduct.

8 Closing Remarks

Once I had completed the data analysis, I tried to acquire more evidence in support of the supposition that the auctioneer owns nicknames and interacts in the auctions. I wrote an email to Mr. Merkli (Managing Director of LABUYLA Switzerland) and explained that my interest in contacting the management of LABUYLA was for the purpose of acquiring information for a scientific paper that I was writing on the auction design of LABUYLA. I also asked whether the company could provide me with a dataset of the website's activities for the purpose of analyzing empirical behavior. A day later, Mr. Perri (Managing Director of LABUYLA Germany) called me by phone and tried to establish exactly what I was investigating in my study and invited me to their office in Stuttgart.

Mr. Perri welcomed me on June 21, 2010 and presented LABUYLA's business plan to me. This presentation included a preview of the new homepage layout, which he said would be online in a few weeks. He said that further adjustments still needed to be made to the auction platform, but that the operating homepage still provided the possibility of buying an item at the online store and at an auction. Mr. Perri regretted that he was unable to give me any data on the activities taking place on the auction platform. He justified this by saying that they had not yet integrated a system for collecting such data and therefore possessed no such information.

I had prepared two graphs of my data for this meeting (details of Figure 4 and Figure 6). I confronted Mr. Perri with these patterns and asked him whether he had any explanation of how these patterns might have occurred. He was unable to give me any reason for why such coordinated behavior occurred. Mr. Lewandowski (Communications Designer) joined us, but could not explain the pattern either. They asked me for my interpretation of how the pattern occurred. I answered that my findings result in only one possibility; namely, that the auctioneer owns nicknames and places bids. The talk stagnated at this point, as Mr. Perri mentioned a few times that such a behavior would not only be illegal, but would not produce any profit. Since he emphatically repeated that such a strategy would not be profitable, I explained how such behavior might indeed be profitable. The discussion was terminated with his admission that they had tried such a strategy in April 2009 (when they launched the auction platform), but that they had lost over CHF 20,000 within a few months, thus proving how unprofitable such a strategy is.

A few weeks after this conversation, LABUYLA launched its new homepage. But instead of the version they had shown me, the new homepage no longer featured the possibility of buying an item in an auction. Since then, LABUYLA has published no explanation for why the items are now only on sale at a fixed price in the online store. They have simply shut down their auction platform, which had generated a profit of over CHF 430,000 in three months, and they would certainly not have closed such lucrative platform if the auctioneer had not owned multiple nicknames.

In its judgment of December 04, 2012 SHAB [2012], the district court of Meilen (Switzerland) declared LABUYLA Holding AG subject to bankruptcy proceedings; and so, the company was dissolved.

A Appendix

A.1 Proof of Proposition 1

Proof. First suppose that $t > 0$ and thus the probability that at least one further noise bidder arrives is positive. $\hat{u}(b) = u(p_t, t, \bar{b}) - u(p_t, t, b)$, where $\bar{b} \equiv b - \Delta$, describes the gain from waiting for an additional noise bidder instead of buying the item at some price b. If $\hat{u}(b)$ is positive, the noise bidder should wait. If $\hat{u}(b)$ is negative, then the noise bidder should either buy the item as soon as the price reaches b or check whether he can do better by choosing a higher b.

$$
\begin{aligned}
\hat{u}(b) =& u(p_t, t, \bar{b}) - u(p_t, t, b) \\
=& \pi_2(n_b + 1, t)\pi_3(n_b + 1, p_t)\left(\bar{p} - b + \Delta\right) \\
& + \sum_{i=n_x}^{n_b} \pi_1(i, t)\pi_3(i, p_t)\left(\bar{p} - p_t + i\Delta\right) \\
& - \pi_2(n_b, t)\pi_3(n_b, p_t)\left(\bar{p} - b\right) - \sum_{i=n_x}^{n_b-1} \pi_1(i, t)\pi_3(i, p_t)\left(\bar{p} - p_t + i\Delta\right) \\
=& \pi_3(n_b, p_t)\left[\pi_2(n_b, t) - \pi_1(n_b, t)\right]\Big\{F(\bar{b})\left(\bar{p} - b + \Delta\right) - \left(\bar{p} - b\right)\Big\} \\
=& \underbrace{\pi_3(n_b, p_t)}_{>0}\underbrace{\left[\pi_2(n_b, t) - \pi_1(n_b, t)\right]}_{>0}\Big\{F(\bar{b})\Delta - \Big[1 - F(\bar{b})\Big]\left(\bar{p} - b\right)\Big\}.
\end{aligned} \tag{3}
$$

The informed bidder's decision depends only on the sign of the term in curly brackets and is independent of the time t and the current price p_t. In order to identify the optimal b, I take a closer look at the sign of the last term. If b is such that this term is less than or equal to zero, then for any $b > b'$ this term is strictly negative, because $F(b - \Delta) > F(b' - \Delta)$ and $(\bar{p} - b) < (\bar{p} - b')$. The sign of the term in curly brackets is either negative for all b or changes at most once from positive to negative. Along with the fact that $\hat{u}(\Delta) < 0$, either $b^* = p_T$, if $\hat{u}(p_T) < 0$, or the optimal price fulfills $\hat{u}(b^*) \leq 0$ and $\hat{u}(b^* + \Delta) > 0$. $\square$

A.2 Proof of Proposition 2

Proof. If $t > 0$ and $x_i > p_T$, then the informed bidder's problem described in Equation (1) can be described as

$$u(p_t, t, b) = \pi_2(n_b, t)\pi_3(n_b, p_t)\,(\bar{p} - b) + \sum_{i=1}^{n_b-1} \pi_1(i, t)\pi_3(i, p_t)\,(\bar{p} - p_t + i\Delta)$$
$$+ \pi_2(n_b, t)\,[1 - \pi_3(n_b, p_t)]\,(x_i - p_T)$$
$$+ \sum_{i=0}^{n_b-1} \pi_1(i, t)\,[1 - \pi_3(i, p_t)] \cdot (x_i - p_T)\,.$$

As before, I calculate the gain if an additional informed bidder arrives.

$$\hat{u}(b) = \pi_3(n_b, p_t)\,[\pi_2(n_b, t) - \pi_1(n_b, t)]\left\{F(\bar{b})\Delta - \left[1 - F(\bar{b})\right](\bar{p} - b)\right\}$$
$$+ \underbrace{\pi_3(n_b, p_t)}_{>0}\underbrace{[\pi_2(n_b, t) - \pi_1(n_b, t)]}_{>0}\underbrace{\left[1 - F(\bar{b})\right]}_{>0}\underbrace{(x_i - p_T)}_{>0}. \tag{4}$$

Equation 4 is strictly larger than Equation 3 and, therefore, the additional option increases his expected profit. □

A.3 Proof of Proposition 3

Proof. The FOC of Equation 2, with $\Omega^*_{z,p_t-\Delta} = \frac{z}{\bar{\omega}}(\bar{p} - p_t + \Delta)$, is

$$\frac{\partial \Omega_{t,p_t}}{\partial \omega_{p_t}} = \underbrace{\frac{e^{-\lambda(t-\omega_{p_t})}}{\bar{\omega}}}_{>0}\underbrace{\left\{(\bar{p} - p_t) + \left[(\bar{p} - p_t) - \pi_3(1, p_t)(\bar{p} - p_t + \Delta)\right]\lambda\omega_{p_t}\right\}}_{\stackrel{!}{=}0}$$

and thus

$$\omega^*_{p_t} = \frac{1}{\lambda}\frac{(\bar{p} - p_t)}{\pi_3(1, p_t)(\bar{p} - p_t + \Delta) - (\bar{p} - p_t)} > 0.$$

This result implies that $\omega^*_{p_t}$ is independent of t and strictly positive. The SOC evaluated at $\omega^*_{p_t}$ is negative, and therefore $\omega^*_{p_t}$ is a maximum. □

A.4 Proof of Proposition 4

Proof. Assume that $p_t \in (\underline{p} + \Delta, \underline{p} + 2\Delta]$. Proposition 2 implies two things: Firstly, $\frac{\partial \omega^*_{p_t}}{\partial p_t} < 0$ and thus $\omega^*_{p_t} < \omega^*_{p_t-\Delta}$; Secondly, $\Omega^*_{z,p_t-\Delta} = \frac{z}{\bar{\omega}}(\bar{p} - p_t + \Delta), \forall z \in (0, \omega^*_{p_t-\Delta})$. Thus

Equation 2 can be written as:

$$
\begin{aligned}
\Omega_{t,p_t} &= e^{-\lambda(t-\omega_{p_t})}\frac{\omega_{p_t}}{\bar{\omega}}(\bar{p}-p_t) \\
&\quad + \int_{\omega^*_{p_t-\Delta}}^{t} \lambda e^{-\lambda(t-z)}\pi_3(1,p_t)\Omega^*_{z,p_t-\Delta}dz \\
&\quad + \int_{\omega_{p_t}}^{\omega^*_{p_t-\Delta}} \lambda e^{-\lambda(\omega^*_{p_t-\Delta}-z)}\pi_3(1,p_t)\left[\frac{z}{\bar{\omega}}(\bar{p}-p_t+\Delta)\right]dz.
\end{aligned} \tag{5}
$$

$\omega^*_{p_t}$ maximizes Equation 2 as well as Equation 5 and thus does not depend on the possibility that the informed bidder might prefer to wait for more than one additional noise bidder for all $p_t \in (\underline{p}+\Delta, \underline{p}+2\Delta]$. This argument can be used iteratively and thus holds for all $p_t \in (\underline{p}, \bar{p})$. □

A.5 Proof of Proposition 5

Proof. Three possible cases exist if $t > \bar{\omega}$.

1. If $\omega^*_{p_t} < \omega^*_{p_t-\Delta} \leq \bar{\omega}$, then only the second term of Equation 5 changes.

2. If $\omega^*_{p_t} \leq \bar{\omega} < \omega^*_{p_t-\Delta}$, then Equation 5 changes to:

$$
\begin{aligned}
\Omega_{t,p_t} &= e^{-\lambda(t-\omega_{p_t})}\frac{\omega_{p_t}}{\bar{\omega}}(\bar{p}-p_t) \\
&\quad + \int_{\omega^*_{p_t-\Delta}}^{t} \lambda e^{-\lambda(t-z)}\pi_3(1,p_t)\Omega^*_{\bar{\omega},p_t-\Delta}dz \\
&\quad + \int_{\bar{\omega}}^{\omega^*_{p_t-\Delta}} \lambda e^{-\lambda(\omega^*_{p_t-\Delta}-z)}\pi_3(1,p_t)(\bar{p}-p_t+\Delta)dz \\
&\quad + \int_{\omega_{p_t}}^{\bar{\omega}} \lambda e^{-\lambda(\bar{\omega}-z)}\pi_3(1,p_t)\left[\frac{z}{\bar{\omega}}(\bar{p}-p_t+\Delta)\right]dz.
\end{aligned}
$$

3. If $\bar{\omega} < \omega^*_{p_t} < \omega^*_{p_t-\Delta}$, then the informed bidder's best action is to play *bid and buy* right away.

In all cases, the maximization still results in the same $\omega^*_{p_t}$ or the same strategy. □

A.6 References to Websites

I often refer to content on labuyla.ch, which (for the reasons explained in this paper) does not exist anymore:

http://www.labuyla.ch/liveAuction/all/, retrieved July 05, 2010;

http://www.labuyla.ch/liveAuction/bestseller/10/, retrieved July 05, 2010;

http://www.labuyla.ch/liveAuction/testAccount/, retrieved July 05, 2010.

Screenshots of these sites are available from me upon request. In addition, I refer to the auction design provided by dealwonders.com and dubLi.com. These two websites do not exist anymore, or no longer offer a price reveal auction:

http://www.dealwonders.com/, retrieved July 10, 2010;

http://global.dubli.com/, retrieved July 10, 2010.

Advertising spots for the price reveal auction which they offered are available (retrieved September 18, 2014) on:

http://www.youtube.com/watch?v=6UbSbkKiNhc, dealwonders.com;

http://www.youtube.com/watch?v=SLHtQcbTyQs, dubLi.com.

Bibliography

ALEEM, A. and ANTWI-BOASIAKO, A. (2011). Internet Auction Fraud: The Evolving Nature of Online Auctions Criminality and the Mitigating Framework to Address the Threat. *International Journal of Law, Crime and Justice*, **39** (3), 140–160.

ARIELY, D., OCKENFELS, A. and ROTH, A. E. (2005). An Experimental Analysis of Ending Rules in Internet Auctions. *RAND Journal of Economics*, **36** (4), 890–907.

AUGENBLICK, N. (2014). *Consumer and Producer Behavior in the Market for Penny Auctions: A Theoretical and Empirical Analysis*. Revise and Resubmit: Review of Economic Studies.

BLOOMFIELD, R., O'HARA, M. and SAAR, G. (2009). How Noise Trading Affects Markets: An Experimental Analysis. *Review of Financial Studies*, **22** (6), 2275–2302.

BYERS, J. W., MITZENMACHER, M. and ZERVAS, G. (2010). Information Asymmetries in Pay-Per-Bid Auctions: How Swoopo Makes Bank. *CoRR*, **abs/1001.0592**.

CHAKRABORTY, I. and KOSMOPOULOU, G. (2004). Auctions With Shill Bidding. *Economic Theory*, **24** (2), 271–287.

CHUA, C. E. H. and WAREHAM, J. (2008). Parasitism and Internet Auction Fraud: An Exploration. *Information and Organization*, **18** (4), 303–333.

DELTAS, G. (1999). When Does Cheating on Mail-In Bids Pay? A Guide for the Dishonest Auctioneer. *Journal of Economic Behavior & Organization*, **40** (3), 313–323.

DONG, F., SHATZ, S. M. and XU, H. (2009). Combating Online In-Auction Fraud: Clues, Techniques and Challenges. *Computer Science Review*, **3** (4), 245–258.

GALLICE, A. (2010). *Price Reveal Auctions on the Internet*. Carlo Alberto Notebooks 147, Collegio Carlo Alberto.

GRAHAM, D. A., MARSHALL, R. C. and RICHARD, J.-F. (1990). Phantom Bidding Against Heterogeneous Bidders. *Economics Letters*, **32** (1), 13–17.

Hossain, T. (2008). Learning by Bidding. *RAND Journal of Economics*, **39** (2), 509–529.

Jenamani, M., Zhong, Y. and Bhargava, B. (2007). Cheating in Online Auction — Towards Explaining the Popularity of English Auction. *Electronic Commerce Research and Applications*, **6** (1), 53–62.

Ockenfels, A., Reiley, D. and Sadrieh, A. (2006). *Online Auctions*. NBER Working Papers 12785, National Bureau of Economic Research, Inc.

— and Roth, A. E. (2006). Late and Multiple Bidding in Second Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction. *Games and Economic Behavior*, **55** (2), 297–320.

Östling, R., Wang, J. T.-y., Chou, E. Y. and Camerer, C. F. (2011). Testing Game Theory in the Field: Swedish LUPI Lottery Games. *American Economic Journal: Microeconomics*, **3** (3), 1–33.

Platt, B. C., Price, J. and Tappen, H. (2010). *Pay-to-Bid Auctions*. NBER Working Papers 15695, National Bureau of Economic Research, Inc.

Rapoport, A., Otsubo, H., Kim, B. and Stein, W. E. (2009). *Unique Bid Auction Games*. Working Paper No. 2009,005, Jena Economic Research Papers.

Robinson, S. G., Giebelhausen, M. D. and Cotte, J. (2013). Shopping, Gambling or Shambling? Penny Auctions. *Journal of Business Research*, **66** (9), 1612–1616.

Ross, S. M. (2010). *Introduction to Probability Models*. Academic Press, 10th edn.

Roth, A. E. and Ockenfels, A. (2002). Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet . *American Economic Review*, **92** (4), 1093–1103.

SHAB (2012). Notice no: 6974050, December 13, 2012, Schweizerisches Handelsamtsblatt, shab.ch.

SNYDER, J. M. (2000). Online Auction Fraud: Are the Auction Houses Doing All They Should or Could to Stop Online Fraud? *Federal Communications Law Journal*, **52** (2), Article 8.

STEIGLITZ, K. (2007). *Snipers, Shills, and Sharks: eBay and Human Behavior*. Princeton University Press.

VGH BADEN-WÜRTTEMBERG (2013). Judgment of May 23, 2013, 6 S 88/13, juris.de.

WANG, R., DINLERSOZ, E. M. and BAG, P. K. (2000). More on Phantom Bidding. *Economic Theory*, **15** (3), 701–707.

YOKOO, M., SAKURAI, Y. and MATSUBARA, S. (2004). The Effect of False-Name Bids in Combinatorial Auctions: New Fraud in Internet Auctions. *Games and Economic Behavior*, **46** (1), 174–188.

Chapter III

Analyzing the Behavior and Influence of Accurately Identified Shill Bidders in Online Auctions

Abstract. By merging data files containing the personal details of bidding agents with millions of bidding histories, this is the first paper that has succeeded in accurately identify shill bidders in online auctions and analyze their behavior. I show that the identified shill bidders influenced 0.35% of the observed auctions. Moreover, I calculate the shill bidders' gains and monetize the auctioneer's incentive to overlook price-pushing shills. The behavior of the majority of these accurately identified shill bidders does not fulfill any of the criteria for the shill bidder types discussed in the literature. Nevertheless, I adopt two algorithms which aim to identify shill bidders based on public information. On average, these approaches assign a higher probability of being a shill bidder to the accounts of the bidders who were accurately identified as shill bidders in my dataset. However, the possibility of reliably identifying shill bidders and honest bidders is limited.

Keywords: online auction; shill bidding; in-auction fraud; bidding behavior

Jelclass: D12; D44

1 Introduction

Shill bidding describes the behavior of a seller who bids in his own auctions with an intention, for example, to drive up the sales price or to avoid the official reserve price's insertion fee. The anonymity of the web facilitates this type of fraudulent activity, because even an attentive observer would hardly be able to assign a user account to a specific individual. "The most crucial issue is that online identities are easily created and cannot be tracked back to the physical identities without inside information" [Ockenfels *et al.*, 2006, p. 25]. The auctioneer, however, has access to the participants' personal details, and is therefore typically in the best position to prevent shill bidding.

The auctioneer's main source of revenue from sold articles is the final value fee. A bonafide warning that shill bidding is prohibited might encourage more aggressive bidding behavior, thereby increasing the auctioneer's profit. Unfortunately, the auctioneer faces conflicting incentives in deciding whether or not to actually fight this fraudulent behavior. On the one hand, he could disregard bidding misconduct in order to personally profit from the price-pushing shill bids. On the other hand, his reputation as a trusted auctioneer is at stake.

Several studies (e.g., Trevathan and Read, 2007; Engelberg and Williams, 2009; Kauffman and Wood, 2005; Ford *et al.*, 2013; Dong *et al.*, 2010) have recognized this incentive problem and have developed approaches, based on bidding histories and further public information, to tackle misconduct of this kind. Their algorithms successfully identify shill bidders' accounts in their training datasets. However, owing to the lack of field data which document proven shill bidders, it is difficult to verify the precision of their suggested approaches.

This paper is based on empirical data which contain almost two million bidding histories, the individuals' maximum bid, as well as the users' personal details (e.g., full postal address and phone number). By comparing the seller's and bidders' personal details in each auction, these data allow me to identify obvious shill bids. I use the term 'obvious' shill bids, because the provided account information is sufficient to identify shill bidder accounts with certainty.

The recorded auctions took place on the online auction platform ricardo.ch. Ricardo.ch AG is a Swiss subsidiary of the multinational media company Naspers. With over 3 million user accounts and 20,000 articles sold per week, ricardo.ch operates the most frequently used auction platform in Switzerland.

Auctions on ricardo, like those on eBay, follow the rules of an ascending English auction and offer the possibility of proxy bidding. An ascending English auction starts at a previously defined reserve price. The auctioneer then raises the price as long as more than one bidder signals his willingness to pay the currently announced price. If only one bidder remains, then this bidder wins the auction at the current price. Such auctions are concluded in a short time, and the bidders' cost of attending the auction is negligible. In contrast, online auctions take place over several days. Bidders do not attend the auction all the time, and returning to the website in order to place a bid is costly in terms of time. The proxy bidding system aims to account for this inconvenience and expense by automatically placing bids up to an entered maximum value.

This paper enhances the existing literature with insights on the behavior of accurately identified shill bidders. I show that current shill bidding has many guises and that identifying the underlying strategies used is difficult. This finding is supported by the limited success of identification algorithms which are based on specific patterns of behavior. In addition, I calculate the influence of obvious shill bids on the rent's distribution. Besides estimating the honest bidders' loss, I focus on the auctioneer's monetary incentive to overlook shill bids, because the final value fee he collects is often used as an argument for the observable persistence of shill bids.

The remainder of the paper is structured as follows. Section 2 discusses the related literature and outlines previous findings. In Section 3, I describe the data collection in more detail and summarize my acquired dataset. Section 4 analyzes the behavior of the accurately identified shill bidders, investigates the success of prohibiting obvious shill bidders, and estimates the auctioneer's monetary gain from shill bids. In Section 5, I apply two shill-bidder identification algorithms to my data in order to check whether they identify the shill bidders my approach accurately identified. As most studies on online auctions use data about behavior on eBay, Section 6 compares the summary statistics of the bidding behavior on ricardo with equivalent figures on eBay (USA) auctions presented by Hayne *et al.* [2003]. The aim of this section is to evaluate to what extend the findings about shill bidding on ricardo can be applied to eBay. Finally, Section 7 concludes.

2 Literature

The rising popularity of online auctions has recently entailed an abundance of studies in this field (Ockenfels *et al.*, 2006; Aleem and Antwi-Boasiako, 2011). One branch of this literature investigates in-auction fraud; i.e., the fraudulent behavior of any participant during a running auction. Such fraudulent behavior can emanate either from the auction house (e.g., Bag *et al.*, 2000), from bidders (e.g., Sher, 2012; Chen and Tauman, 2006), or from the seller (e.g., Chakraborty and Kosmopoulou, 2004; Izmalkov, 2004; Hlasny, 2006). The latter contains shill bidding.

2.1 Motivating Shills

Shill bidding describes the fraudulent in-auction behavior of a single individual, who influences the auction by using multiple accounts in order to submit bids (Yokoo *et al.*, 2004). Even though auctions exists where potential buyers shill bid to hide their purpose (Sher, 2012; Yokoo *et al.*, 2004), I use the term 'shill bid' in this narrower sense and focus only on sellers who bid on articles they are selling. Shills can differ with regard to the sellers' motivation and actions (e.g., Kaur and Verma, 2013; Kauffman and Wood, 2003; Engelberg and Williams, 2009; Watanabe and Yamato, 2008; Izmalkov, 2004): (1) Reserve-Price Shilling occurs early in the auction and ensures a minimum price, without paying the fees associated with an official reserve price; (2) Buy-Back Shilling occurs late in the auction and is aimed at prohibiting the sale of the article; (3) with Competitive Shilling, the seller continuously outbids the highest bidder (up to a certain amount); (4) Discover-and-Stop Shilling (which I describe below); (5) creating phony feedback to inflate reputation; or, for example, (6) exposing the bidders' valuation for a subsequent offer.

The Discover-and-Stop strategy is based on eBay's proxy bidding rules (Engelberg and Williams, 2009). Suppose the current highest bidder H entered b_H as his maximum bid and the minimum increment is Δ. Later on, another bidder A enters the amount b_A as his maximum bid. Therefore, four cases are possible: (1) if $b_H \geq b_A + \Delta$, then the standing price increases to $b_A+\Delta$, and H is still the highest bidder; (2) if $b_A+\Delta > b_H \geq b_A$, then the standing price increases to b_H, and H is still the highest bidder; (3) if $b_H + \Delta > b_A > b_H$, then the standing price increases to b_A, and A becomes the highest bidder; or, (4) if $b_A \geq b_H + \Delta$, then the standing price increases to $b_H + \Delta$, and A becomes the highest bidder. Whenever the second case is observed (H remains the highest bidder, but the standing price increases to less than $b_A + \Delta$), then b_H has been discovered and the shill bidder stops placing shills.

2.2 Models Which Include Shill Bidding

Chakraborty and Kosmopoulou [2004] analyze shill bidding incentives in a common value auction setting where the probability that the seller can participate in his own auctions is exogenous and common knowledge. They show that bidders cannot be fooled in equilibrium, because they correctly anticipate the seller's potential participation and, therefore, bid less aggressively. Since the equilibrium strategies imply that the seller sometimes wins an auction, the possibility of shill bidding lowers the expected payoff for buyers and sellers. In contrast, the auction house, which only cares about the sales price and not about who wins the auction, gains from shill bidding.

Chakraborty and Kosmopoulou [2004] mention that the toleration of shill bidders potentially affects the auctioneer's long-run profitability due to reputation losses. However, they do not consider reputation effects in their model. Kosmopoulou and Silva [2007] provide experimental evidence for this model. In line with the model's findings, they show that bidders bid less aggressively if they believe that shill bidders are present.

Izmalkov [2004] investigates the implications of a shilling seller who sells a single object to asymmetric buyers in a private values setting by means of an English auction. His main finding is that "there exists an equilibrium of the English auction with shill bidding that is outcome equivalent to the optimal mechanism of Myerson [1981]" [Izmalkov, 2004, p. 1]. In addition, he discusses several issues, such as resale possibilities, which should be considered in real-life applications.

2.3 Evidence of Shill Bidding in Online Auctions

Although shill bidding is strictly prohibited, court rulings (Schwartz and Dobrzynski, 2001) prove that this behavior distorts the course of numerous online auctions. Furthermore, the anonymity provided by internet auctions facilitates shilling and is additionally advantaged when the auction house permits multiple accounts per person. Auction houses have an important role to play in preventing shill bids. Along with bidding histories, they have exclusive access to the personal data of all the accounts. Therefore, the auctioneer should be able to link accounts and detect obvious shill bidders.

The fact that auction houses might have insufficient incentives to fight shill bidding has prompted several studies which attempt to identify shill bidders based on observable bidding behavior. These studies estimate that the share of auctions that are influenced by shill bidders potentially ranges between 1% and 10% (Ockenfels and Roth, 2006). However, this range must be treated with caution, because some of these studies (e.g., Engelberg and Williams, 2009) base this share only on auctions with more than one bidder, while other studies refer to 'questionable bids' instead of shill bids (e.g., Kauffman and Wood, 2003).

2.4 Attempts to Identify Shill Bidders on the Basis of Bidding Histories

In this section, I will briefly outline the approaches taken by four studies to identify shill bidders. Without access to personal details, these authors use the bidders' observable behavior as a proxy for shilling. These studies suggest 'shilling variables' which capture suspicious bidding characteristics and describe how to calculate and condense them into a 'shill score' or into a shill probability, respectively. They test their approaches either on simulated or empirical auction data and successfully identify shill bidders according to their specific strategy.

- Trevathan and Read [2005] developed an algorithm which calculates a shill score based on six shilling variables. This shill score indicates whether a seller and a specific buyer are engaged in collusive shill-biding behavior. In Trevathan and Read [2007], they extend their approach in order to identify sellers who use multiple accounts to place shill bids. Their suggested collusion score combines the shill score with dependencies among nicknames using a graph theory approach.

- Engelberg and Williams [2009] show that eBay's proxy bidding design allows for a Discover-and-Stop strategy. By incremental bidding, a shill bidder can at times discover the maximum bid of the highest bidder and stop bidding just below this value. Based on bidding behavior characteristics, they estimate that 1.39% of all bids on eBay are placed according to their identified Discover-and-Stop strategy and are therefore shills. Engelberg and Williams [2009] use a probit model to validate their suggested shilling variables. Moreover, they experimentally demonstrate the profitability of the Discover-and-Stop strategy and highlight the associated risk of accidentally becoming the highest bidder. They became the highest bidder in 16 out of 46 Discover-and-Stop attempts and won 7 out of 30 auctions.

- Kauffman and Wood [2005] use premium bids as a proxy for shills. A premium bid is a bid which is placed at an auction while another auction exists on the same article at a lower current price. They find that premium bidding occurs in 23% of all auctions in their dataset.

- Ford *et al.* [2013] developed an algorithm based on a feed-forward, back-propagation neural network in order to detect suspicious bidders. Suspicious bidders can then be analyzed in more detail by an external shill verifier. Such a shill verifier is, for example, suggested by Dong *et al.* [2010] who use an approach based on the Dempster-Shafer theory of evidence.

3 Data

In 2013, Naspers owned three additional European auction houses besides ricardo.ch, namely: ricardo.gr (operating in Greece, closed on May 5th, 2014); qxl.dk (operating in Denmark, sold in the autumn of 2013); and qxl.no (operating in Norway, sold in the autumn of 2013). Along with ricardo.ch, these four auction houses formed the ricardo

Group. As their auction platforms all used the same software, all of them unintentionally submitted the highest bidder's maximum bid in each auction to the browser of each observer along with the auction's bidding history. In addition for ricardo.ch and ricardo.gr, the personal details of every account were publicly available. This unique opportunity to accurately identify obvious shill bidders owing to the accessibility of personal details is the reason for my focus on auctions that took place on ricardo.

3.1 Origin of the Data

In order to display dynamically changing contents in a user's browser, the ricardo Group's websites use 'Asynchronous JavaScript and XML' (AJAX) requests. In practice, this means that whenever ricardo.ch is addressed, then: (1) the browser receives a source code which contains placeholders; (2) the browser requests the current values for those placeholders from the server; (3) the server replies those values; and (4) the browser combines all information in order to display the requested website.

In December 2013, I noticed that very valuable information was submitted by ricardo.ch's server in response to the implemented AJAX requests. Along with the public variables' values (e.g., nickname), the server submitted personal details such as the user's postal address or his (mobile) phone number. Whenever an individual was browsing ricardo.ch's auctions, the browser received the personal details of all participating users. Even though this confidential information was not displayed in the browser, its submission clearly breaches data protection rules. Privacy protection concerning the account data is not only in the consumers' interest. Ricardo itself is anxious to prevent any communication between auction participants, as the participants may collude and arrange price agreements in order to avoid auction fees.

Shill bidders who attempt to raise the sales price, but also wish to avoid ending up as the highest bidder, typically have to make decisions under uncertainty about the maximum bid that another (proxy) bidder has entered. However, the AJAX standard request for the bidding history of (currently running) auctions on ricardo returned the value of the entered maximum bid along with the bidding history. Therefore, the combination of ricardo's proxy bidding system and its information leakage which submits the highest bidder's maximum bid is virtually an invitation to shill.

On April 22, 2014, ricardo.ch blocked my IP address owing to the high traffic my intense activity caused and, therefore, I stopped the recording. I advised ricardo.ch of the information leakage on their platform and they closed it within a few hours. After repeated requests, qxl.dk and qxl.no closed the leakage on their website in late July 2014.

3.2 Data Collection

The data recording occurred in the period between December 21, 2013 and April 22, 2014. At the end of an auction, its bidding history is typically available for the next two months. This allowed me to also record auctions which had ended prior to the recording period. With regard to the total value of sales price, this recorded dataset approximately captures a quarter of the yearly sales on ricardo.ch.

The requests that were made to ricardo.ch's server in order to gain the desired items of information are described in the Appendix. Similar requests were possible for ricardo.gr, qxl.dk and qxl.no. Only the domain and the value of the variable 'PartnerNr' had to be changed according to Table 1.

Ricardo.ch does not prohibit to use computer systems for gathering information about the activities undertaken on their platform, as long as these requests do not interfere with its normal performance. Collecting and using these (personal) data for scientific research conforms with the Swiss data protection law, DSG, Article $13^2(e)$.

For the analyses and figures in this paper, I disregard all recorded auctions where either the item was sold by the buy-it-now option, multiple-items were sold in a single auction, or an item was offered by the auction house for test purposes. In addition, I study only manually entered (maximum) bids and disregard consecutively placed proxy bids by the system.

3.3 Dataset Summary

Before focusing on shill bidding, I briefly discuss some key figures in the recorded dataset shown in Table 1. These values refer to sold articles and, therefore, to auctions with at least one bidder. The number of bids equates the number of times a bidder entered a new (maximum) bid. In order to compare the number of accounts across platforms, ‘active’ accounts refer to accounts which were involved in at least one recorded auction as a bidder or seller.

Table 1: Dataset Summary

Domain	ricardo.ch	ricardo.gr	qxl.dk	qxl.no
PartnerNr	2	14	12	20
Number of auctions	1,776,459	5,302	90,383	90,389
Number of bids	7,013,419	11,209	261,584	242,575
Number of accounts	3,053,462	516,122	NA	NA
Number of active[1] accounts	496,083	3,854	17,796	12,533
Average bids per auction	3.95	2.11	2.89	2.68
Average bidders per auction	2.49	1.59	2.04	1.91
Average page views per auction	137	75	31	29
Median page views per auction	70	31	22	17
Auctions without reserve price (%)	34%	24%	13%	9%

Notes: This table summarizes the recorded dataset.

[1] Active means, that the account was involved in at least one recorded auction as a bidder or as a seller.

Ricardo.ch is by far the most frequently used platform. In comparison with the other ricardo Group platforms, over twenty times as many articles were sold via the Swiss auction house. In addition, auctions on ricardo.ch attract on average the most bids and the most bidders. The number of bidders and bids refers to the number of observed, not potential, bidders and bids in an auction. Because bidders do not attend the auction constantly throughout its duration, the order in which they arrive at the auction as well as their strategy influences these observed values. Suppose, for example, that all bidders bid their true valuation and all valuations differ by more than the auction-specific minimal increment. If they arrive in ascending order of their valuations, then all bidders and their bids are observed. If, however, the two bidders with the highest valuations arrive first, then only their bids are observed, because the remaining bidders’ valuations are below the current minimum bid.

The optimal reserve price in a standard auction model exceeds the seller's valuation and does not depend on the number of bidders involved (e.g., Krishna, 2009). The auction house encourages sellers to abstain from a reserve price, which is the same as the starting price on ricardo, by giving a listing priority to auctions which give with a starting price of CHF 1 and by charging insertion fees which typically increase with higher reservation prices. The data show that up to one third of the sellers abstained from setting a reserve price. These sellers either attach nearly no value to their item and, therefore, the provided benefits of no-reserve price exceed the sellers' valuation, or they did not set an optimal reserve price.

The final value fee on ricardo.ch was not observable and has to be calculated. During the observation period this fee had a staggered structure. For the first CHF 200 of the sales price, the final value fee was 6.5%; for the next CHF 800, the final value fee was 4.5%, and amounts beyond CHF 1,000 were charged for 2.5%. If, for example, an item has been sold for CHF 1,400, then the final value fee was CHF 59 ($200 \times 0.065 + 800 \times 0.045 + 400 \times 0.025$). The final value fee differs for items listed in the category 'motors', including 'parts and accessories'. The final value fee for these items amounted CHF 29, CHF 39, CHF 49 or CHF 59, depending whether the sales price was in the range CHF 0–1,000, CHF 1,000.01–3,000, CHF 3,000.01–5,000 or above CHF 5,000.01.

4 Shill Bidding on ricardo

Owning multiple accounts is not prohibited in the terms and conditions of ricardo. However, ricardo explicitly prohibits shill bids placed either: (1) by the seller; (2) by a person living in the seller's household; or, (3) by any person acting on behalf of the seller. Breach of this paragraph is sanctioned by issuing a caution or suspension.

4.1 Identifying Obvious Shill Bidders on the Basis of Personal Data

The accounts' personal details enable me to identify obvious shill bidders. Accounts which are registered with identical full name, address, postal code and zip are assigned to a single person (*fullAddress*). Identical mobile phone numbers also indicate the same person and the identical landline phone numbers indicate that the persons live in the same household (*phoneNumber*). In addition, I treat accounts with identical last names, addresses, postal codes and zip codes as individuals inhabiting the same household (*shortAddress*), being aware of the fact that this criterion might contain incorrect relationships.

Table 2 shows that on ricardo.ch more than half of all the accounts belong to individuals who own one single account. Most individuals with multiple accounts are identified through their phone number, whereas half of all users specified their mobile phone number.

Table 2: Identification of Multiple Accounts

	ricardo.ch		ricardo.gr	
	Accounts	Individuals	Accounts	Individuals
phoneNumber	813,720	348,296	299,885	134,635
fullAddress	378,506	171,102	33,666	15,507
shortAddress	726,383	305,270	42,202	19,161
potentialShillBidAccounts	1,413,590	487,970	329,891	140,252

Notes: This table shows: (1) the number of individuals who own multiple accounts; (2) the number of accounts these individuals own; and, (3) through which personal details variable the assignment occurs. As some accounts might be identified through more than one of these criteria, *potentialShillBidAccounts* reports the number of accounts that fulfill at least one criterion.

This approach does not have the capacity to identify relationships between individuals who have different last names or who do not live at the same address. In particular, I cannot assign two accounts to one specific individual if the person moved and opened up a new account afterwards.

A shill bidder cannot be identified simply by the number of accounts an individual owns, but instead by whether or not he uses more than one account in a specific auction. Table 3 shows the number of identified shill bid accounts as well as the number of individuals owning them. Of the half a million potential shill bidding individuals, only two thousand are accurately identified to be shill bidders.

On ricardo.gr, the share of auctions in which a shill bid occurred is 0.15% of the

Table 3: Identified Shill Bid Accounts

	ricardo.ch		ricardo.gr	
	Accounts	Individuals	Accounts	Individuals
identicalUser	none		none	
phoneNumber	2,518	1,214	6	3
fullAddress	284	142	none	
shortAddress	1,950	967	2	1
shillBidAccounts	3,902	1,895	8	4

Notes: This table shows: (1) the number of individuals who are accurately identified shill bidders; (2) the number of accounts these individuals used for their shills (seller and buyer accounts); and, (3) through which criterion the they are identified. As some accounts might be identified through more than one of these criteria, *shillBidAccounts* reports the number of accounts that fulfill at least one criterion.

total. However, this share results from only eight influenced auctions. For this reason, the following shill bid analysis addresses solely the shill bidding behavior on ricardo.ch.

The 3,902 *shillBidAccounts* on ricardo.ch placed shill bids in 6,183 auctions (0.35% of the total). If only auctions with more than one bidder are considered, then the share of influenced auctions almost doubles to 0.66%. These values are lower than previous findings in the literature. However, they solely represent a lower bound on shill bidding because shill bidders who use, for example, third-party accounts are not identified. Auctions with the minimum starting price of CHF 1 were affected more than three times as much from shill bids than auctions with a reserve price.

Regarding product categories, auctions listed in the category ‘motors’ were influenced most often, with 1.8%. The correlation between a category’s mean selling price and the share of influenced auctions is 0.80. Shill bidders tend to bid in categories where high-valued items are sold and where, therefore, a percentage increase in the final sales price results in correspondingly high profits.

The data show that the correlation between the mean number of honest bidders and the share of influenced auctions in a category is 0.89. Auctions on motors, for example, attract on average 6.3 honest bidders, which is more than twice as many as found for the other categories. Moreover, auctions which were influenced by shill bids show on average a higher participation rate of honest bidders in all categories. This result is contrary to my expectation; especially, for auctions with a low number of bidders or weak competition I would have expected to observe relatively more price-driving shill bids.

4.2 Profiteers and Victims of Shill Bids

Since the final value fee rises with the sales price, the auctioneer gains from shill bids. This positive relationship is often used to explain for the lack of incentive that auctioneers have in countering shill bidding. In this subsection, I determine the auctioneer's additional profit which arises from shill bids in order to gain fresh insight on the strength of this argument. I further evaluate the seller's fraudulently gained surplus, as well as the honest bidders' losses.

I calculate these values by eliminating shill bids from auction histories, while holding all other bidding decisions constant. This approach ignores the fact that (non-observed) bidders would probably have placed further bids as a result of a lower standing bid, because including such effects is not possible without major assumptions. For the same reason, I further abstain from including strategic interaction (e.g., Chakraborty and Kosmopoulou, 2004) or optimal decisions such as the choice of the starting price.

The redistribution effects depend on whether the shill bidder won the auction or was the second highest bidder.

- In 1,944 auctions, neither of the above was the case and, therefore, the final sales price has not been influenced by shill bids.

- 1,677 auctions ended with the shill bid as the second highest bid. These shill bids increased the total sales price by approximately CHF 67,000 and increased the auctioneer's revenue from final value fees by CHF 2,000 (2.8% of the redistributed buyer rents). On average, these bidders bought the items at a price that was CHF 40, or 25%, higher than it would have been without shill bids.

- Shill bidders won 2,562 auctions with a total value of CHF 600,000. For these auctions, the auction house charged them final value fees of over CHF 21,000.

The large share of unsuccessful shill bids might indicate that not all of them had the intention of actually selling the item. Even though I can only speculate about their intentions, incentives to win their own auction might arise, for example, from: (1) increasing the accounts reputation through the rating system; (2) learning about the item's market value for a later sale; (3) avoiding an obligation to sell the item at a low price as a result of weak competition and a suboptimal reserve price; or, (4) direct benefits from such fictional sales.

In summary, honest bidders were cheated for the amount of CHF 67,000, which corresponds to 0.05% of the total sales value. This fraudulently gained amount is accompanied by final value fees of successful and unsuccessful shills amounting to CHF 23,000. These payments to the auctioneer account for 0.27% of all the observed auctions' final value fees.

4.3 Ricardo's Success in Prohibiting Obvious Shill Bidding

Identifying shill bids through personal details is the easiest way of detecting the obvious case of shill bidding. Such behavior could be automatically detected in real-time, without requiring much effort from the auction house. I do not claim to judge ricardo's attempt to fight other more sophisticated and, potentially, more important shill bidding strategies. Moreover, I do not have any data on the resources spent on preventing fraud in general. However, the presence in the auction of bidders who can repeatedly place shill bids is not in line with a successful enforcement of ricardo's terms and conditions.

Table 4 displays such repeated shill bidding activities. During my observation period, shill bids were placed by 1,895 different individuals. In total, 54% of them shilled in one single auction. Either these individuals were cautioned and abstained from further shill bids, or the observation period did not capture further shills. However, more important than this percentage, are the remaining 46%.

According to ricardo's terms and conditions, individuals who place bids for reasons other than winning the auction get cautioned for their first offense. For a second offense, they are barred from the platform. Therefore, all individuals who influenced at least two auctions through shill bidding should have been barred from further auction participation,

Table 4: Influence of Single Shill Bidders

influenced auctions	number of individuals	involved accounts	blocked accounts	suspended accounts
1	1,016	2,032	67	47
2	335	694	28	11
3-20	507	1,084	60	18
21-40	25	63	5	0
41-60	5	14	1	1
61-80	3	6	0	0
over 80	4	9	2	0
Total	1,895	3,902	163	77

Notes: This table shows how many auctions a single bidder has influenced through a shill bid account. The number of blocked or suspended accounts refer to the accounts' status on April 22, 2014.

and their accounts should have been closed. But only 7% of the accounts involved in shill bidding activities were either suspended or closed on April 22, 2014. The analyses in this section show that shill bidding with a second user account, containing identifying personal details, was not consequentially and efficiently sanctioned by ricardo.ch.

The fact that the two blocked accounts that are responsible for manipulating over 80 auctions do not belong to the same individual implies that these accounts must have been blocked due to other offenses. If these accounts would have been blocked due to their shill bidding activities, then the second account of these individuals would have been blocked as well.

4.4 Categorizing Identified Shill Bids

The literature generally describes four different types of shill bidding, namely: Reserve-Price Shilling; Competitive Shilling; Buy-Back Shilling; and Discover-and-Stop Shilling. The aim of this subsection is to determine which shilling strategy most commonly occurs on ricardo.ch. Shill bids are, as in the previous subsections, identified either through the *phoneNumber*, the *fullAddress*, or the *shortAddress*. Rather than categorizing the account, I regard and categorize each shill bid or each shill bid sequence, respectively. This approach takes into account that the majority of the individuals use different shill bidding strategies.

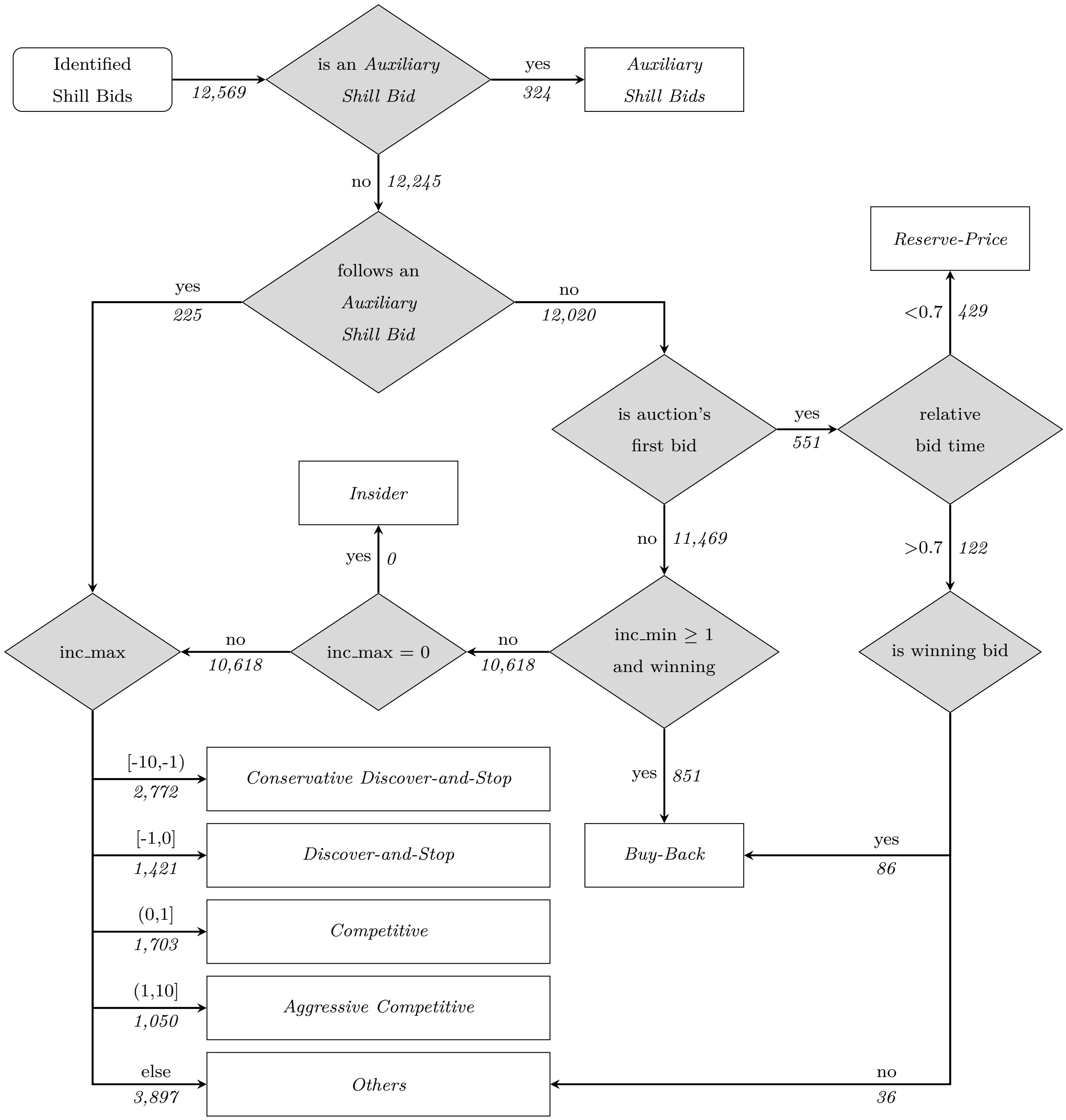
The Discover-and-Stop strategy is not applicable on ricardo. If $b_H \geq b_A$, then ricardo's proxy bidding system places a bet at b_A for H. H remains the highest bidder, because earlier bids are given priority. However, I assign bids as if this strategy would be possible on ricardo in order to examine whether the behavior of some shill bidders nonetheless meets the Discover-and-Stop Shilling criteria.

The Competitive Shilling strategy as well as the Discover-and-Stop Shilling strategy typically need more than one bid. I name these bids 'Auxiliary Shill Bids'. As entered maximum bids can be publicly observed on ricardo, a shill bidder can simply bid the same maximum price in order to raise the current sales price. I call this strategy 'Insider Shilling'.

In order to increase the number of bids that can be assigned to a specific bid type, I define two additional types which allow for a more tolerant interpretation of the Discover-and-Stop and Competitive Shilling strategy. The Conservative-Discover-and-Stop Shilling strategy allows the bidder to underbid the maximum bid by up to 10 increments and considers that the expected profit loss from overbidding the highest bidder increases along with higher prices. This risk causes the shill bidder to stop discovering the highest price, even though the hidden maximum price is still more than one increment over the current highest bid. Another interpretation would be, that the shill bidder simply tries to raise the price, while trying to avoid becoming the highest bidder. The Aggressive-Competitive Shilling strategy allows a shill bidder to overbid the current maximum bid by up to 10 increments. This strategy reduces the shill bidder's number of bids that he has to submit for a specific article. The time that is thereby saved might offset the higher risk of buying the article in the end, especially if the competing bidder does not use the proxy bidding system. If a shill bid does not satisfy the criteria of any of these eight types, then this bid is assigned to the type 'Others'.

Figure 1 shows the criteria according to which I categorize identified shill bids into

Figure 1: Categorizing Identified Shill Bids



Notes: This figure shows how I categorize identified shill bids into different shill bidding types. The decisive criteria are:

is an *Auxiliary Shill Bid*:	If the shill bid is followed by another shill bid within 10 minutes and one proxy bid occurred between these two shill bids.
follows an *Auxiliary Shill Bid*:	This shill follows — but is not itself — an *Auxiliary Shill Bid*.
is auction's first bid:	If this bid is the first bid of the auction.
relative bid time:	Bid time relative to the auction's duration.
is winning bid:	If this bid is the winning bid of the auction.
inc_min:	The bid's increment compared to the minimal bid required.
inc_max:	The bid's increment compared to the current *maxBid*.

The italic numbers express how many bids feature a specific characteristic.

different shill bidding types as well as how many bids feature a specific characteristic. Around one fifth (22%) of the 12,569 shill bids are identified as (Aggressive) Competitive Shill bids, and about one third (33%) of the shill bids follow the (Conservative) Discover-and-Stop Shilling strategy. Insider Shilling was never observed. This shows that the information leakage about the current maximum bid remained unexploited by the identified shill bidders.

This section's findings imply that the reliable identification of shill bidders based on publicly observable behavior is a difficult task. Around one third (31%) of the identified shill bids do not match any of the labled strategy's criterion and are categorized as 'Others'. Moreover, shilling individuals were identified as having pursued multiple shilling strategies during the observed period.

5 Shill Bidder Identification Algorithms

In this subsection, I adopt two algorithms, one suggested by Trevathan and Read [2005] and the other by Engelberg and Williams [2009], and I check whether their suggested methods can reliably identify the proven shill bidder accounts in my dataset.

5.1 Algorithm Suggested by Trevathan and Read [2005]

Trevathan and Read [2005] identify shill bidders based on the weighted average of six shilling variables, namely:

α: The participation rate of the bidder in auctions provided by a particular seller.

β: The bidder's average percentage of bids submitted throughout all the auctions he participated in.

γ: The share of auctions the bidder won (normalized value).

δ: The bidder's average inter-bid time (normalized value).

ϵ: The bidder's average inter-bid increment (normalized value).

ζ: The average time of the bidder's first submitted bid (normalized value).

The bidders' behavior is summarized in a shill score: $SS = \frac{w_1\alpha+w_2\beta+w_3\gamma+w_4\delta+w_5\epsilon+w_6\zeta}{w_1+w_2+w_3+w_4+w_5+w_6} \times 10$. All of the shilling variables' values range between 0 and 1. A higher value implies a higher probability that this bidder is shilling. Decisive for the calculation of β, δ and ϵ are the entered maximum bids and not the bids placed by the proxy bidding system. If a bidder won the auction, his α, β, δ, ϵ and ζ values for this auction are set to zero. Further details on the shilling values' calculation and their associated weights are provided in Trevathan and Read [2005].

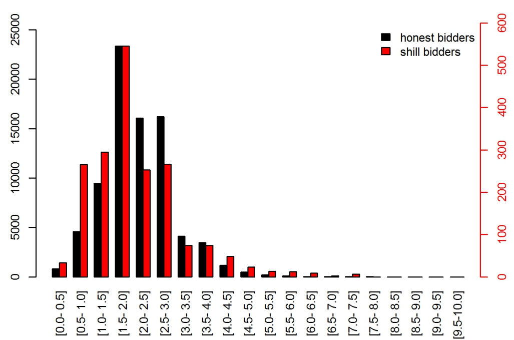
I check their method on a subset for which the algorithm should perform best. This subset consists auctions offered by accurately identified shilling individuals. I further exclude 33 sellers who used more than one shill bidder account in a specific auction and another 79 sellers who used more than one shill bidder over all auctions. Therefore, if an individual is not marked as a shill bidder, then this individual is almost certainly an honest, non-shilling bidder. Finally, the analyzed subset contains 1,913 sellers who sold 49,650 articles. In this analysis, 81,921 pairs of sellers and bidders are rated.

For each pair of sellers and bidders, I calculate: the shill score as well as the shill score ranking of a bidder compared to the other bidders who bid on articles offered by this seller. Figure 2 shows, on the left-hand side, the score's or ranking's distribution for shill bidders as well as honest bidders. The Student's t-test as well as the Kolmogorov-Smirnov-test are significant to the 1 percent level and therefore imply that the types' score distributions differ. The right-hand side of Figure 2 shows the algorithm's false-positive rate and false-negative rate for different threshold values.

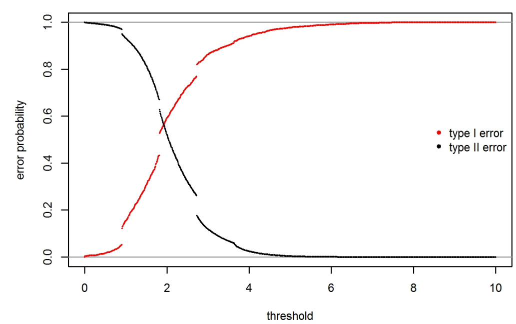
Even though the shill score and the ranking provide an indication of whether the bidder's behavior is suspicious, the type II error prevents a reliable identification of the proven shill bidders in this subset. For example: If bidders whose shill score ranking equals 1 are incriminated as shill bidders, then 223 (11.6%) of the actual shill bidders will be correctly incriminated, while 1,851 (2.3%) of the honest bidders will be falsely incriminated.

Figure 2: Performance of the Trevathan and Read [2005] Identification Algorithm

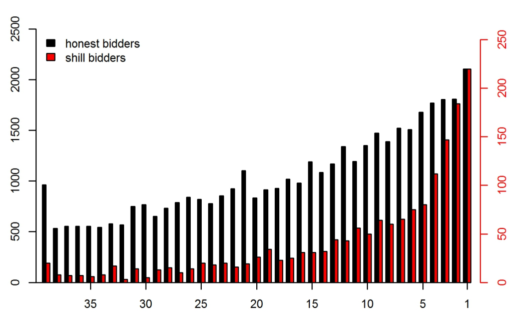
(a) Shill Score: Distribution



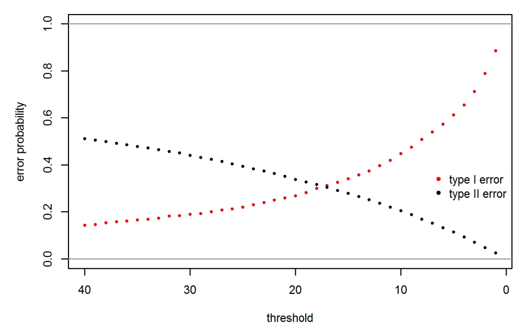
(b) Shill Score: Identification Errors



(c) Shill Ranking: Distribution



(d) Shill Ranking: Identification Errors



Notes: The histograms on the left-hand side show the score's or, respectively, the ranking's distribution for shill bidders and honest bidders. The figures on the right-hand side show the type I and type II error of the bidder's identification for different threshold values.

5.2 Algorithm Suggested by Engelberg and Williams [2009]

Engelberg and Williams [2009] identify shill bidders by using their newly identified Discover-and-Stop strategy as a proxy for shill bidding. They estimate a probit regression model that is based on six publicly observable variables in order to identify bidders who follow the Discover-and-Stop strategy. These shilling variables are:

$BidderCount$: The number of bidders in an auction.

$ClosingPrice$: The sales price at which the auction ended.

$BRating$: The user rating as a proxy for experience.

$NumAuctions$: The number of auctions provided by a particular seller in which the bidder participated.

$FracBid$: The participation rate of the bidder in auctions provided by a particular seller.

$FracLose$: The share of auctions the bidder lost.

In order to separate the calculation of $NumAuctions$, $FracBid$ and $FracLose$ from the Discover-and-Stop identification, Engelberg and Williams [2009] split their sample into two distinct time periods. Engelberg and Williams [2009] considered only Event Ticket auctions, because they expected heterogeneous private values and considerable bid amount dispersion in this category.

I am able to use my whole observation period to calculate the variable values, because shill bidders are identified according to their personal details and not according to their behavior. Table 5 shows the probit regressions' output based on several subsets of my ricardo.ch dataset.

The subsets used for the regression results in Column (1) and Column (2) contain auctions from all categories. In contrast to Column (1), Column (2) is based on auctions which were solely offered by shilling sellers. The coefficient values in Column (3) refer to a subset which includes only Event Ticket auctions.

Column (4) compares most closely to the subsets used by Engelberg and Williams [2009] which are shown in Column (5) and Column (6). The coefficients of the regression

Table 5: Performance of the Engelberg and Williams [2009] Identification Algorithm

	ricardo.ch				eBay[1]	
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	-3.732***	-3.196***	-3.807***	-40.8	-2.093***	-2.158***
	(0.013)	(0.035)	(0.174)	(147.6)	(0.021)	(1.078)
BidderCount	-0.013***	-0.024***	-0.048***	-0.396***	0.021***	0.0022**
	(0.001)	(0.002)	(0.014)	(0.131)	(0.003)	(0.009)
ClosingPrice	32.2***	73.7***	-430.9	-6,952	-4.40	-10.50
	(4.0)	(6.9)	(390)	(6,442)	(21.3)	(16.6)
BRating	-0.074***	-0.287***	-0.854***	-0.023	-0.3691***	-0.2145
	(0.007)	(0.014)	(0.205)	(0.240)	(0.061)	(0.163)
NumAuctions	-2.33***	-2.27***	24.8***	59.3***	6.52**	7.71***
	(0.12)	(0.14)	(3.57)	(12.8)	(2.03)	(2.09)
FracBid	1.478***	1.640***	1.577***	38.7	0.221***	0.350***
	(0.013)	(0.034)	(0.166)	(147.6)	(0.051)	(0.071)
FracLose	0.883***	1.106***	1.185***	1.860***		-0.033
	(0.014)	(0.020)	(0.128)	(0.663)		(0.090)
# of observations	4,415,037	157,459	51,400	35,337	104,545	10,087
# of shill bids observations	6,293	6,291	119	15	2,711	295
Only auctions from shilling seller	No	Yes	No	No	No	No
Only DiscNStop shills	No	No	No	Yes	Yes	Yes
Only losing (B,S) pairs	No	No	No	Yes	Yes	Yes
Only Event Tickets	No	No	Yes	Yes	Yes	Yes

Notes: Standard errors are in parentheses. *, ** and *** indicate significance at the 10 percent, 5 percent and 1 percent level, respectively. Closing Price is multiplied by 1,000,000, BRating is multiplied by 1,000, and NumAuctions is multiplied by 1,000 to reduce the number of decimal places.

[1] Comparable figures on eBay auctions originate from Engelberg and Williams [2009].

output differ with regard to their value and significance. Compared to Column (5) and Column (6), the coefficient values for BRating and FracBid are not significant for the ricardo sample, FracLose is significant and BidderCount shows a negative coefficient value. For this subset, the algorithm is able to reliably identify shill bidders; i.e., reliable in terms of correctly incriminating 73% of actual shill bidders, while only falsely incriminating 44 (0.1%) of the honest bidders. However, my subsample includes only 15 shill bidders who used a strategy that looks like Discover-and-Stop shilling.

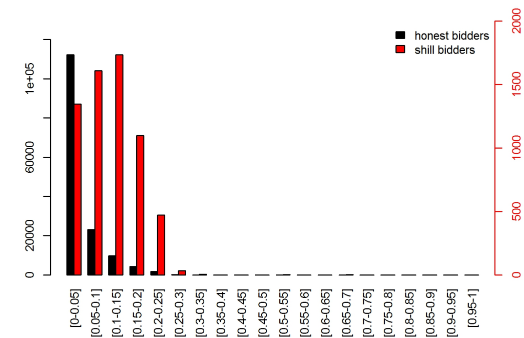
The positive coefficient for FracLose is in line with the argument that shill bidders do not usually intend to win the auction. The reverse but still significant value for BidderCount can be explained by the fact that a higher number of bidders increase competition and lead to a more aggressive bidding behavior. In such an environment, shill bids are not as profitable as they are in auctions with only a few bidders.

The regressions' coefficient values differ with regard to their significance as well as whether they increase or decrease the bidder's probability of being a shill bidder. The differences in these regression outputs might be explained by the fact that the Discover-and-Stop Shilling strategy cannot be used on ricardo and, therefore, these shill bidders' behavior just looks like Discover-and-Stop Shilling.

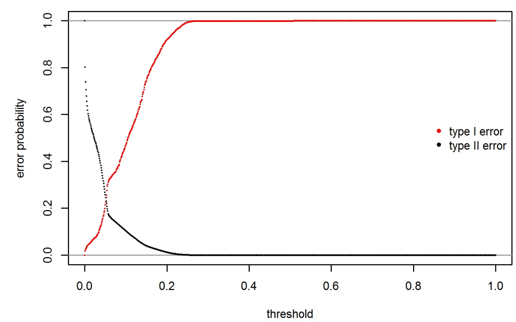
Figure 3 shows the distribution of the estimated probability that a bidder shills as well as the type I and type II error of the bidder's identification. The distribution of the estimated probability might indicate that the algorithm suggested by Engelberg and Williams [2009] can reliably identify shill bidders in other subsets than the one used for Column (4). However, the axis scaling absorbs the large imbalance between honest and shill bidders. The type II error reveals that achieving a low type I error without incriminating too many honest bidders of shilling is not possible.

Figure 3: Performance of the Engelberg and Williams [2009] Identification Algorithm

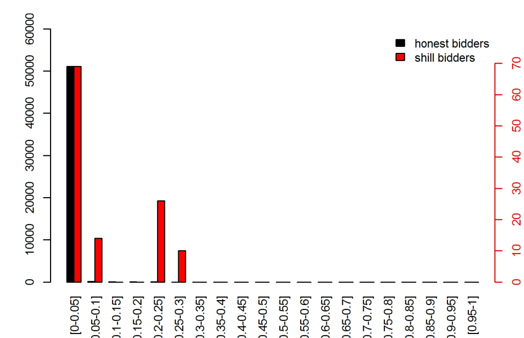
(a) Distribution: Table 5, Column (2)



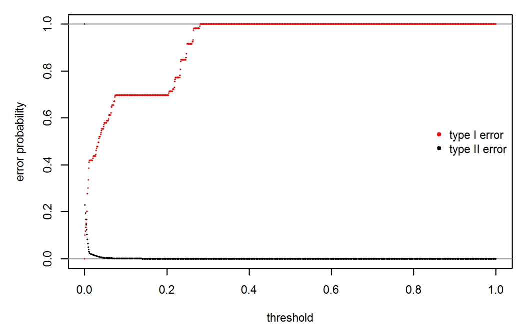
(b) Identification Errors: Table 5, Column (2)



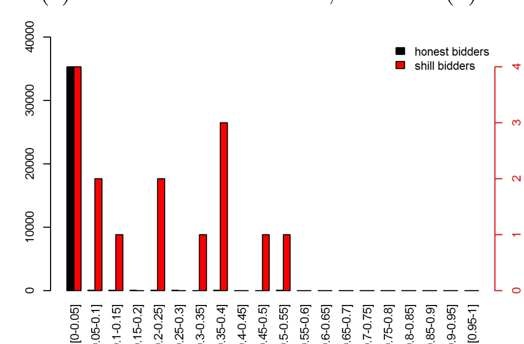
(c) Distribution: Table 5, Column (3)



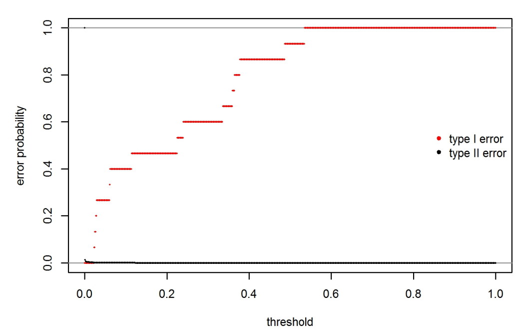
(d) Identification Errors: Table 5, Column (3)



(e) Distribution: Table 5, Column (4)



(f) Identification Errors: Table 5, Column (4)



Notes: The histograms on the left-hand side show the distribution of the estimated probability that a bidder shills, for both; shill bidders and honest bidders. The figures on the right-hand side show the type I and type II error of the bidder's identification for different threshold values.

6 Analyzing Bidders' Behavior

This section compares selected characteristics of the bidders' behavior across eBay and the four platforms provided by the ricardo Group. The comparative figures for 11,495 U.S. based auctions on eBay originate from Hayne *et al.* [2003]. The transnational comparison in Table 6 aims to assess whether or not the previous results on ricardo.ch can be applied to other platforms. In this section, I use only data from auctions with at least two bidders. On the one hand, these auctions are more interesting, as they allow observations, such as multiple bids per bidder in an auction. On the other hand, since Hayne *et al.* [2003] calculated his values for auctions with more than one bidder, I adopt this restriction.

The term 'unique bidders' refers to the number of different, active accounts in the dataset. In contrast, the total number of bidders sums up the number of unique bidders per auction over all auctions. Only one tenth of all auctions on ricardo.gr received bids from more than one bidder. In contrast to this, on ricardo.ch, qxl.dk, and qxl.no one half of all auctions had more than one bidder. Table 6 shows that each individual (unique bidder) recorded in ricardo's data placed around 10 bids on average, whereas only a few individuals in the data provided by Hayne *et al.* [2003] placed more than one bid. The average experience of the active bidders is very different across the platforms. On the one hand, the average bidders' experience for ricardo.gr is much lower compared to the other platforms, which is not surprising considering the number of articles sold. On the other hand, qxl.dk and qxl.no have a five-times higher average bidders' experience than ricardo.ch and eBay. However, these averages are driven by outliers.

The mean number of bids and of unique bidders is similar across all platforms. My dataset contains the information about the winning bidder's maximum bid that he entered in the proxy bidding system. Therefore, all bids, or respectively bidders, can be classified for ricardo. In addition, Table 6 shows that on average the winner's entered proxy bid is between 20% and 50% higher than the mean sales price. The values of the articles sold on the Norwegian platform are on average substantially higher than for those sold on other platforms.

Table 6 shows that the five platforms are similar regarding the bids' and the bidders'

Table 6: Transnational Comparison of Bidders' Behavior in Auctions

	ricardo.ch	ricardo.gr	qxl.dk	qxl.no	eBay[1]
Data Summary					
number of auctions,	906,336	1,143	38,716	37,778	11,495
total number of bids	9,325,449	9,578	265,663	243,647	77,926
total number of unique bidders	386,834	1,744	11,824	10,062	40,754
total number of bidders	3,595,801	4,437	132,563	119,776	45,797
accounts' average experience rating	127	29	830	694	112
mean number of bids	10.3	8.4	6.9	6.5	6.8
mean number of (unique) bidders	4.0	3.9	3.4	3.2	4.0
mean starting price	42.4	28.8	40.9	96.1	19.4
mean winning bid amount	110.5	53.4	123.7	251.6	63.1
winner's mean entered proxy bid	166.2	56.3	150.9	322.60	NA
Bids Classified by Frequency					
single bid (%)	22.5	28.6	31.0	30.0	38.6
multi-bid (%)	77.5	71.4	69.0	70.0	61.4
Auction Success of Bidders Classified by Frequency					
single bid bidders (%)	58.3	61.7	62.2	61.1	65.7
auctions won (%)	50.6	55.5	53.8	55.0	57.9
success rate (%)	55.1	60.4	58.8	60.7	NA
average experience rating	990	100	2,819	3,142	155
multi-bid bidders (%)	41.7	38.3	37.8	38.9	34.3
auctions won (%)	49.4	44.5	46.2	45.0	42.1
success rate (%)	67.8	68.0	67.6	64.7	NA
average experience rating	547	82	2,664	2,352	100
Bids Classified by Bid Strategy					
proxy bids (%)	37.3	20.8	49.4	51.8	75.0
incremental bids (%)	62.7	79.2	50.6	48.2	20.1
unclassifiable bids (%)	0	0	0	0	4.9
Auction Success of Bidders Classified by Bid Strategy					
only proxy bidders (%)	19.7	11.2	20.6	22.9	73.1
auctions won (%)	46.6	27.6	47.9	53.8	72.1
success rate (%)	62.0	50.8	69.3	72.3	81.2
average experience rating	829	232	2,160	2,515	145
only incremental bidders (%)	33.0	57.6	24.9	21.8	13.6
auctions won (%)	26.3	55.7	19.3	14.7	5.6
success rate (%)	32.8	58.8	24.6	20.9	12.6
average experience rating	698	58	3,259	3,483	133
proxy and incremental bidders (%)	47.4	31.2	54.5	55.3	10.1
auctions won (%)	27.1	16.6	32.8	31.5	9.5
success rate (%)	52.6	50.8	56.8	54.2	29.4
average experience rating	534	90	2,669	2,353	78
unclassifiable bidders (%)	0	0	0	0	3.2

Notes: This table shows a transnational comparison of bidder's behavior in auctions with more than one bidder.

[1] The comparative figures for 11,495 U.S. based auctions on eBay originate from Hayne *et al.* [2003].

frequency classification. In contrast, the four platforms of the ricardo Group differ from eBay with regard to the use of the proxy bidding system. Ricardo as well as eBay recommend proxy bidding to the bidders (entering their article's valuation), as this behavior is the weakly dominant strategy in a second-price sealed bid auction. One reason for the observation of a higher share of incremental bidders who place multiple bids per auction might be the different methods for identifying proxy bids. My dataset includes a variable which indicates proxy bids. In contrast, an identification based on the bidding time falsely identifies a fast series of incremental bids as proxy bids.

Incremental bidding is often (e.g., Ockenfels and Roth [2006]) associated with inexperienced bidders who behave as if the articles are auctioned in a first-price auction without proxy bidding. However, my dataset does not support the supposition. Incremental bidders are usually just as experienced as proxy bidders and bidders who place both, proxy and incremental bids.

Ockenfels and Roth [2006] analyze how different ending rules influence bidder's behavior and might explain the observed behavioral differences. Ricardo uses a 'soft close' ending rule with an automatic run-time expansion of three minutes, whereas the run-time on eBay does not depend on the bids' timing ('hard close'). If bidders place proxy bids which are lower than their valuation (or if they adjust their valuation over time), then the time expansion rule on ricardo might tempt bidders to react to other bidders' late bids. In contrast to this, eBay's hard close ending rule favors sniping, which does not allow for further (incremental) bids and results in the higher success rate of proxy bids on eBay.

Even though Table 6 reveals some differences between the platforms, on the whole the five platforms show quite similar values regarding the bidders' behavior.

7 Conclusion

Confidential information such as an individual's personal details and the maximum bid entered into the proxy bidding system is not publicly available and is generally subject to data protection. By exploiting an information leakage, I was able to record this information along with the bidding histories of almost two million auctions which took place on ricardo.ch.

By using the accounts' personal details, this is the first study that does not need to infer shill bidders from a specific kind of behavior. In contrast to the existing literature, these auction data provide a unique opportunity to accurately identify obvious shill bidders. Moreover, I am able to calculate the redistribution effects of shilling, to assess ricardo's success in preventing obvious shill bidding, and to test shill-bidder identification algorithms.

I provide evidence that at least 0.35% of all auctions on ricardo.ch were influenced by obvious shill bids. This value is a lower bound on actual shill bidding, because more cautious shill bidders, who, for example, use third-party accounts to place bids remain unidentified.

Approximately two thousand individuals placed obvious shill bids in 6,183 auctions. In 1,677 of these auctions, the auction ended with the shill bid being the second highest bid. These shill bids increased the final sales price on average by CHF 40, or 25% respectively.

Repeated obvious shill bidding was not consequentially and efficiently sanctioned on ricardo.ch. However, the additional final value fees from shill bids compared to the total final value fees earned (0.27%) is very low. This finding invites an analytical reappraisal of the auctioneer's incentive to overlook shill bids. I doubt that the auctioneer intentionally allows shills in order to slightly increase his profit. To me it seems more likely, that either this fraudulent behavior was treated as a bearable loss or the possibility to identify obvious shill bids through personal details has not yet been recognized and implemented.

According to shill-bidding types which have been characterized in previous studies, Competitive Shilling, where the seller continuously outbids the highest bidder up to a certain amount, occurred most often. However, around one quarter of the identified shill bids did not match any type's criterion and this share would have been considerably higher without the newly introduced strategies 'Conservative-Discover-and-Stop Shilling' and 'Aggressive-Competitive Shilling'.

I adopt two methods which aim to identify shill bidders based on public information in order to check whether the proven shill bidders can be reliably identified according to the suggested algorithms. Both algorithms, one suggested by Trevathan and Read [2005] and the other by Engelberg and Williams [2009], assign on average a higher probability of being a shill bidder to actual shill bidders. Using a subsample from my data that corresponds to the criteria of the dataset used by Engelberg and Williams [2009], I am able to show that their algorithm reliably identifies shill bidders based on public information. Notwithstanding this case, both algorithms incriminate a greater number of truthful bidders than actual shill bidders who are identified. Although these approaches may indicate suspicious accounts, their type II errors are too high for a reliably identification.

The auction house is in an optimal position to prohibit (obvious) shill bids, and, thereby, customers should expect the auction house to invest substantial effort in enforcing its terms and conditions. Even though several honest customers were cheated by obvious shill bidders, this paper shows that this is a rather rare event and that the auctioneer's monetary incentive to overlook shill bids is relatively low. The strategies of these shill bidders are diverse and often do not even match any type characterized in the literature. With this finding in mind, the limited success of identifying shill bidders based on bidders' publicly observable behavior is not surprising. The transnational comparison of bidder's behavior shows that bidding behavior is almost the same over all platforms provided by the ricardo Group and eBay (USA). Therefore, the results of this paper should be generally applicable to other auction platforms.

8 Appendix

Structure of the the requests to ricardo.ch's server in order to gain the desired information.

User Details Ricardo assigns a unique user number *userNr* to each user account. This identification number remains the same, irrespective of whether the personal details or the *nickName* change.

The requests took the following structure:

http://www.ricardo.ch/DataService/Proxy.aspx?DataService.svc/json/
GetAuctionSellerInfos?UserNr=*userNr*&PartnerNr=2&IsHttps=false

The server's response to the request on my account contained, amongst other things, the following excerpt:

[...] 'PhoneNumber':'0041615546909' [...] 'City':'Basel' [...] 'Street':'Winkelriedplatz','StreetNr':'9','ZipCode':'4053' [...] 'UserFirstName':'Dominic', 'UserLastName':'Herzog' [...]

Table 7 lists the variables which were recorded for each account and indicates for which platform the corresponding value was submitted. On ricardo.ch and ricardo.gr, the algorithm requested and recorded the personal details of all the respective 3,054,104 and 516,206 currently available user accounts.

Table 7: Submitted Variables Upon a Server Request on User Details

Variable Name	Description	ricardo. ch	ricardo. gr	qxl. dk	qxl. no
userNr	unique user ID	x	x	x	x
nickName	user nickname	x	x	x	x
* *firstName*	user first name	x	x		
* *lastName*	user last name	x	x		
* *street*	user address: street	x	x		
* *streetNr*	user address: street number	x	x		
* *compStreet*	user address: street complementary	x	x		
* *zip*	user address: ZIP code	x	x		
* *city*	user address: city	x	x		
* *country*	user address: country	x	x		
* *phoneNr*	user (mobile) phone number	x	x		
* *company*	company's name	x	x		
memberSince	registration date	x	x	x	x
rating	positive ratings minus negative ratings received	x	x	x	x
sharePosRating	share of positive ratings received	x	x	x	x
boughtItems	total number of items bought	x	x	x	x
soldItems	total number of items sold	x	x	x	x
pos_2	positive ratings received during the last 2 months	x	x	x	x
pos_6	positive ratings received during the last 6 months	x	x	x	x
pos_12	positive ratings received during the last 12 months	x	x	x	x
pos_all	overall positive ratings received	x	x	x	x
pos_diffUser	*pos_all* received from heterogeneous users	x	x	x	x
neu_2	neutral ratings received during the last 2 months	x	x	x	x
neu_6	neutral ratings received during the last 6 months	x	x	x	x
neu_12	neutral ratings received during the last 12 months	x	x	x	x
neu_all	overall neutral ratings received	x	x	x	x
neu_diffUser	*neu_all* received from heterogeneous users	x	x	x	x
neg_2	negative ratings received during the last 2 months	x	x	x	x
neg_6	negative ratings received during the last 6 months	x	x	x	x
neg_12	negative ratings received during the last 12 months	x	x	x	x
neg_all	overall negative ratings received	x	x	x	x
neg_diffUser	*neg_all* received from heterogeneous users	x	x	x	x
* *buyerStatus*	values: 0-4 (interpretation unknown)	x	x	x	x
* *hasGlasses*	values: TRUE and FALSE (interpretation unknown)	x	x	x	x
* *isCertified*	values: TRUE and FALSE (interpretation unknown)	x	x	x	x
* *memberClass*	values: 0-4 (interpretation unknown)	x	x	x	x
* *sellerStatus*	values: 0-4 (interpretation unknown)	x	x	x	x
* *sellerType*	values: 0-4 (interpretation unknown)	x	x	x	x
* *stats*	values: 0-640 (interpretation unknown)	x	x	x	x
* *userRating*	rating the user gave	x	x	x	x
suspended	account has been blocked or suspended	x	x		
suspendedDate	date when the account has been blocked or suspended	x	x		

Notes: This table lists and describes the variables which were recorded for each account. Unfortunately, the interpretation of seven variables could not be accurately revealed. The variables' values were submitted and recorded for the platforms marked with an 'x'.

Auction Details and Their Bidding Histories Ricardo assigns a unique *articleID* to each auction. This number remains the same even if an article has not been sold and the item is re-auctioned. The algorithm stored *articleID*s from current auctions and requested the auction's information once the auction was closed. The requests took the following structure:

http://auto.ricardo.ch/DataService/Proxy.aspx?DataService.svc/json/
GetBidsHistory?AuctionNr=*articleID*&PartnerNr=2&NbBidsToShow=20000
&IsArchived=false&IsHttps=false

The auction's details and the bidding history were recorded for each of the two million auctions. Table 8 and Table 9 list the recorded variables.

Table 8: Submitted Variables Upon a Server Request on Auction Details

Variable Name	Description
articleID	referred article ID
requestDate	request date
startPrice	auction's starting price
fixPrice	auction's fixed-price
sellPrice	final sales price
* *maxBidPrice*	the winner's maximum bid entered
bidIncrement	minimal bid increment
it_highestBidder	winner's nickname
highest_bidder_id	winner's user ID
endDateLong	auction's end date
* *startDateLong*	auction's start date
timeLeftLong	time left until the auction ends
seller_id	seller's user ID
it_sellerNick	seller's nickname
it_condition	article's condition
it_availability	number of items available
it_pageViews	number of page requests
it_delivery_link	delivery expenses
it_ArticleDescription	article's description

Notes: This table lists and describes the variables which were recorded for each auction. All variables' values were submitted and recorded for ricardo.ch, ricardo.gr, qxl.dk and qxl.no.

Table 9: Submitted Variables Upon a Server Request on the Auction History

Variable Name	Description
articleID	referred article ID
userNr	unique user ID
nickName	user nickname
BidPrice	bid
* *MaxBidPrice*	maximum bid
BidDate	bid's date
BidStatus	values: $0, 1, 4, 8$
BiddedQuantity	number of items desired

Notes: This table lists and describes the variables which were recorded for each auction. The *userNr* has not been recorded for qxl.dk and qxl.no. All other variables' values were submitted and recorded for ricardo.ch, ricardo.gr, qxl.dk and qxl.no. The values for the variable *BidStatus* mean: (0) this bidder placed a higher bid later on; (1) winning bid; (4) retrieved bid; and, (8) this bidder's highest bid.

Bibliography

Aleem, A. and Antwi-Boasiako, A. (2011). Internet Auction Fraud: The Evolving Nature of Online Auctions Criminality and the Mitigating Framework to Address the Threat. *International Journal of Law, Crime and Justice*, **39** (3), 140 – 160.

Bag, P. K., Dinlersoz, E. M. and Wang, R. (2000). More on Phantom Bidding. *Economic Theory*, **15** (3), 701–707.

Chakraborty, I. and Kosmopoulou, G. (2004). Auctions With Shill Bidding. *Economic Theory*, **24** (2), 271–287.

Chen, C.-L. and Tauman, Y. (2006). Collusion in One-Shot Second-Price Auctions. *Economic Theory*, **28** (1), 145–172.

Dong, F., Shatz, S. M. and Xu, H. (2010). Reasoning Under Uncertainty for Shill Detection in Online Auctions Using Dempster-Shafer Theory. *International Journal of Software Engineering and Knowledge Engineering*, **20** (07), 943–973.

Engelberg, J. and Williams, J. (2009). eBay's Proxy Bidding: A License to shill. *Journal of Economic Behavior & Organization*, **72** (1), 509–526.

Ford, B. J., Xu, H. and Valova, I. (2013). A Real-Time Self-Adaptive Classifier for Identifying Suspicious Bidders in Online Auctions. *The Computer Journal*, **56** (5), 646–663.

Hayne, S. C., Smith, C. and Vijayasarathy, L. R. (2003). Who Wins on eBay: An Analysis of Bidders and Their Bid Behaviours. *Electronic Markets*, **13** (4), 282–293.

Hlasny, V. (2006). Testing for the Occurrence of Shill-Bidding (in Internet Auctions). *Graduate Journal of Social Science*, **3** (1), 61–81.

Izmalkov, S. (2004). *Shill Bidding and Optimal Auctions*. Working Paper, Department of Economics, MIT, Cambridge, MA.

KAUFFMAN, R. J. and WOOD, C. A. (2003). Running Up the Bid: Detecting, Predicting, and Preventing Reserve Price Shilling in Online Auctions. In *Proceedings of the 5th International Conference on Electronic Commerce*, ICEC '03, New York, NY, USA: ACM, pp. 259–265.

— and — (2005). The Effects of Shilling on Final Bid Prices in Online Auctions. *Electronic Commerce Research and Applications*, **4**, 2005.

KAUR, A. and VERMA, G. (2013). Shill Detection Techniques and Research Challenges in Online Auction. *International Journal of Advanced Research in Computer Science and Software Engeneering*, **3** (10), 1303–1308.

KOSMOPOULOU, G. and SILVA, D. G. D. (2007). The Effect of Shill Bidding Upon Prices: Experimental Evidence. *International Journal of Industrial Organization*, **25** (2), 291 – 313.

KRISHNA, V. (2009). *Auction Theory*. Academic Press.

MYERSON, R. B. (1981). Optimal Auction Design. *Mathematics of Operations Research*, **6** (1), 58–73.

OCKENFELS, A., REILEY, D. and SADRIEH, A. (2006). *Online Auctions*. NBER Working Papers 12785, National Bureau of Economic Research, Inc.

— and ROTH, A. E. (2006). Late and Multiple Bidding in Second Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction. *Games and Economic Behavior*, **55** (2), 297 – 320, Mini Special Issue: Electronic Market Design .

SCHWARTZ, J. and DOBRZYNSKI, J. H. (2001). 3 Men Are Charged With Fraud In 1,100 Art Auctions on eBay. *The New York Times*, March 9, 2001.

SHER, I. (2012). Optimal Shill Bidding in the VCG Mechanism. *Economic Theory*, **50** (2), 341–387.

Trevathan, J. and Read, W. (2005). *Detecting Shill Bidding in Online English Auctions*. Working Paper, James Cook University.

— and — (2007). Investigating Shill Bidding Behaviour Involving Colluding Bidders. *Journal of Computers*, **2** (10).

Watanabe, T. and Yamato, T. (2008). A Choice of Auction Format in Seller Cheating: A Signaling Game Analysis. *Economic Theory*, **36** (1), 57–80.

Yokoo, M., Sakurai, Y. and Matsubara, S. (2004). The Effect of False-Name Bids in Combinatorial Auctions: New Fraud in Internet Auctions. *Games and Economic Behavior*, **46** (1), 174 – 188.