Bidding Behavior, Seller Strategies, And The Utilization of Information in Auctions For Complex Goods

Inaugural-Dissertation

zur Erlangung des Grades
Doctor oeconomiae publicae (Dr. oec. publ.)
an der Ludwig-Maximilians-Universität München
Volkswirtschaftliche Fakultät

2009

vorgelegt von

Arno Robin Schmöller

Erstgutachter:	Prof. Dr. Klaus M. Schmidt
Zweitgutachter:	Prof. Dr. Martin Kocher
Datum der mündlichen Prüfung:	18. Januar 2010
Promotionsabschlussberatung:	10. Februar 2010

To my parents

Acknowledgements

First and foremost I want to thank my thesis supervisor Klaus M. Schmidt. He was not only most helpful in inspiringly discussing ideas, insightfully commenting early drafts of my papers, and encouraging me, but also willingly wrote numerous reference letters for scholarships and summer schools.

I am also deeply indebted to Martin Kocher for agreeing to serve as second supervisor on my committee and for the insightful comments and suggestions he provided on numerous occasions. Ray Rees completes my thesis committee as third examiner, which I gratefully acknowledge, as well as his encouragement, helpful comments, and the numerous reference letters he provided for me. I also want to thank him for always being welcome at his chair during the time of my doctorate.

I am particularly grateful to my co-author Florian Englmaier for working as a team in three research projects, which are all part of this thesis. His economic thinking and creativity are outstanding. I am grateful to him for being always available for vivid discussions and insightful comments and for supporting and encouraging me throughout the last four years.

During the course of writing this dissertation, I enjoyed inspiring discussions with many people. I have benefited greatly from the professional help and insightful feedback of Philippe Aghion, David Laibson, Ulrike Malmendier, Sander Onderstal, Matthew Rabin, Sven Rady, Monika Schnitzer, Stefan Trautmann, and Joachim Winter. Furthermore, I am indebted to my colleagues Tobias Böhm, Georg Gebhardt, Florian Heiß, Susanne Hoffmann, Karolina Kaiser, Joachim Klein, and Anton Vasilev who provided a lot of impulse and feedback in countless discussions. From the remaining faculty I want to thank Jan Bender, Rene Cyranek, Matthias Fahn, Nico Klein, Sandra Ludwig, Elisabeth Meyer, Julius Pahlke, Felix Reinshagen, Ludwig Reßner, Linda Rousova, Nicolas Sauter, Christina Straßmair, Martin Watzinger, and Hans Zenger for their valuable input. Matthias Dischinger deserves a special place for all those vivid debates and his encouragement, motivation, and helpful comments and suggestions. I was lucky to have colleagues whom I also count to my closest friends.

It has been an honor to belong to the Munich Graduate School of Economics and I also gratefully acknowledge the financial support granted by the German Research Foundation (DFG) that supported my research in various ways and enabled me to attend several international conferences and workshops.

I also owe gratitude to Silke Englmaier, Brigitte Gebhard, and Irmgard von der Herberg for their indispensable and hearty help in handling all kinds of administrative issues.

For the individual papers I owe gratitude to numerous people.

For *Does Bidding for Complex Goods Reflect All Relevant Information? Field Evidence From Online Gaming*, the second chapter of this dissertation, which is joint work with Florian Englmaier, we are indebted to Hattrick Ltd. for their cooperation and giving us the opportunity to conduct this project. We are especially obliged to Ron Brandes for his great support, interest, and valuable hints and to Siegfried Müller for helping us to establish this contact. We thank Jan Bender, Tobias Böhm, Matthias Dischinger, Florian Heiß, Ulrike Malmendier, Klaus Schmidt, Stefan Trautmann, Joachim Winter, and seminar participants at the Munich Research Workshop "Empirical Economics", the Annual Meeting of the EEA 2008 in Milan, and the conference Economics and Psychology of Football 2008 in Innsbruck for their comments and suggestions. Very special thanks go to Hans Zenger for inspiration.

For *Determinants and Effects of Reserve Prices in Hattrick Auctions*, the third chapter of this dissertation, which is joint work with Florian Englmaier, we thank Tobias Böhm, Matthias Dischinger, Stefan Trautmann, and the participants at the Research Workshop "Empirical Economics" in Munich and at the IMEBE 2009 in Granada, and two anonymous referees for their helpful comments and suggestions. Ines Helm provided excellent research assistance. An earlier version of this paper is circulated as CESifo Working Paper No. 2374 under the title *Reserve Price Formation in Online Auctions*.

For *The Evaluation of Complex Goods - Evidence From Online Car Sales*, the fourth chapter of this dissertation, which is also joint work with Florian Englmaier, we thank Matthias Dischinger and Klaus Schmidt for their helpful comments and suggestions. We are especially grateful to Anton Vasilev for his great support with the dataset and his valuable feedback.

For *Strategic Seller Actions in Auctions with Asymmetric Bidders*, the fifth chapter of this dissertation, I am deeply indebted to Florian Englmaier, Ray Rees, Klaus Schmidt, and Monika Schnitzer. Furthermore I benefited from comments by Tobias Böhm, Matthias Dischinger, Susanne Hoffmann, Elisabeth Meyer, and the participants at the Theory Workshop in at the University of Munich.

Most importantly, I am deeply indebted to my parents, Katharina and Günther Schmöller, my family, and my friends for their inexhaustible love and support, their motivation and trust in my decisions.

Arno R. Schmöller
Munich, September 2009

Contents

Chapter 1

Introduction

In recent years, auctions have become increasingly important as a way to determine the price for items on sale. Governments use auctions to assign contracts and privatize state-owned assets, and through the rise of the internet also millions of households found themselves exposed to the challenges of a novel environment for economic activities. With the recent proliferation of online market platforms such as *amazon.com* or *eBay.com*, sales on the internet have become increasingly popular and constitute a real alternative to the classic retail business, and many of these platforms employ some auction format as a way to allocate goods among their customers.

The theoretical literature on the economics of auctions that has emerged since the seminal contributions of Vickrey (1961), Riley and Samuleson (1981), Myerson (1981), and others, gives us a fundamental understanding of how a rational subject should optimally behave in different auction environments. Today, the field covers a wide range of topics ranging from the optimal selling and bidding strategies over multiple-unit auctions to collusion among bidders and bidding rings to name only a few. In general, the most basic task for bidders is to find an optimal bidding strategy according to their valuation for the item on sale that, conditional on the employed mechanism, ensures them a maximum rent in case they win the auction. Similarly, auctioneers are challenged to make choices that ensure them a maximum expected payoff from the auction, e.g. by implementing the optimal auction format or, if the latter is predetermined, setting a revenue maximizing reserve price for a given mechanism.

However, auctions are still an important and active field of research. Many of the theoretical predictions rely on a set of simplifying assumptions which are not always met in practice. In particular, the frameworks of auction markets are complex and potentially affected by numerous confounding factors. By now a broad range of empirical studies has documented and established substantial deviations from fully rational behavior in many areas of individual decision making. In the context of auctions on the internet, economists have analyzed different auction formats under differing information regimes (e.g. Lucking-Reiley, 1999), phenomena like last minute bidding (e.g. Roth and Ockenfels, 2002), or the existence of a winner's curse (e.g. Bajari and Hortacsu, 2003).[1] Regarding a micro-foundation of the determinants of bidder and seller behavior, however, with few exceptions (e.g. Reiley, 2006, Lucking-Reiley et al., 2007, Lee and Malmendier, 2007) the empirical evidence is rather scarce. Importantly, other than in financial markets, in auctions it are potentially those people who make the biggest mistakes that determine the final price. In light of this fact and the increased popularity of auctions on the internet, it is important to understand what the choice sets of buyers and sellers are, and what motivates and drives their decisions in practice.

A second crucial development for this analysis is that many of the goods that are auctioned off have become increasingly complex in nature to the extent that they consist of a multitude of characteristics, and thus of multiple dimensions of quality. With increasing frequency, bidders are challenged to evaluate goods like mobile phones, personal computers, or even cars on the basis of a plethora of information provided on the various attributes when forming their bids. The same applies to the sellers, when facing the task to choose their reserve price, which in turn requires thorough deliberations on the expected valuations of the potential bidders in the population. While economic theory suggests that a rational agent should pick all relevant pieces of information that are available on such goods, there is only little empirical work done along these lines.

In this doctoral thesis I analyze the determinants of bidder and seller behavior in (online) auctions and similar environments to gain new insights into the efficiency of information aggregation and the interplay of various factors that influence peoples' choices. In doing so, I exploit the enhanced availability of large data sets on auctions and sales on the internet, which constitute an unique testing ground to empirically analyze whether the theoretical predictions are in accordance with the actual behavior observed in the field.

[1] For a comprehensive survey on the existing empirical literature on auctions on the internet see e.g. Bajari and Hortacsu (2004) and Lucking-Reiley (2000).

In particular, Chapter 2 empirically analyzes bidding behavior in auctions within the highly controlled virtual economy of a popular online game and provides strong evidence for a systematic underutilization of available information. Closely related to this study, Chapter 4 examines whether a similar effect persists in a real-world market that involves large stake purchase decisions and further substantiates the presence of biased information processing. The remaining two chapters are concerned with the instruments of sellers to influence the auction outcome. Extending the analysis from Chapter 2 to the supply-side of the virtual-economy, Chapter 3 deals with the determinants and effects of reserve prices and the implications of suboptimal minimum bids in terms of expected revenue. In the context of auctions with asymmetric bidders, Chapter 5 studies theoretically the possibilities and incentives of a seller to favor a specific bidder by using other instruments than reserve prices and direct mechanisms of discrimination. Each chapter is self-contained and can be read independently.

Chapter 2 of this thesis is based on the paper DOES BIDDING FOR COMPLEX GOODS REFLECT ALL RELEVANT INFORMATION? FIELD EVIDENCE FROM ONLINE GAMING.[2] In this paper, we take a novel approach to empirically examine individual bidding behavior in a highly controlled auction environment, where all crucial information on the offered complex goods is openly accessible. Using detailed field data from HATTRICK, one of the largest and most popular online games, we analyze which of the characteristics of a virtual football player, essentially a multi-dimensional vector of attributes, are reflected in the winning bids when he is traded in an auction on the games internal transfer market. We present strong evidence that bidders systematically fail to efficiently aggregate the information provided on the age attribute of these players and thus inefficiently utilize crucial parts of valuable information when forming their bids. The users in the game seem to disproportionately cling to the figure displayed in the nosier information on the age-group of a player, while being inattentive to the finer information embodied in the precise age in units of days. As a consequence, the pattern of winning bids exhibits substantial discontinuities for otherwise identical players, i.e. close substitutes, as many people systematically overpay for players close to their birthday. Since the market environment in the game is highly structured and provides a considerable degree of control, some of the potential sources commonly brought forward in the light of overbidding in other auction markets do not apply. For instance, the

[2]This paper is joint work with Florian Englmaier.

effect is too systematic to be compatible with a non-standard utility of winning ("bidding fever") and also framing effects due to heterogeneous product descriptions can be ruled out. Moreover, by exploiting the natural experiment of a small change in the game design that reflects an exogenous variation of the search cost in the market framework, we are able to analyze whether this "birthday effect" can be explained by classic search costs. We find that the intensity of the documented frictions due to inefficient information processing is clearly attenuated but still preeminent, suggesting that the observed individual behavior is at least partly a result of heuristic decision making. However, these findings clearly illustrate that available information is not always used efficiently and that seemingly minor details of the search environment may have a substantial impact on the outcome of auctions.

In Chapter 3, which is based on the paper DETERMINANTS AND EFFECTS OF RESERVE PRICES IN HATTRICK AUCTIONS,[3] we use a different hand-collected data set of HATTRICK auctions to extend the analysis from Chapter 2 to the supply side of this virtual-economy. In particular, we study the determinants and effects of reserve prices. Conveniently, unlike many other online auction platforms (e.g. *ebay.com*), in this framework there is no proportional relation between the minimum bid and the transaction fees a seller is charged, which could bias the individuals in their choice of a reserve price. Moreover, for the duration of the auction all relevant information concerning the quality of a virtual player becomes publicly available, such that there is no information asymmetry between buyers and sellers and the reserve price is the latters' only variable of choice. It turns out that we find evidence for both, very sophisticated and suboptimal behavior by the sellers. On the one hand, reserve prices are adjusted remarkably nuanced to the resulting sales price pattern. This reflects the theoretical prediction that the sellers take into account the expected valuations of the potential bidders when choosing their reserve price. On the other hand, we provide evidence for the sunk cost fallacy as there is a substantial positive effect on the reserve price when the player has been acquired previously, even though the market environment is highly competitive. In addition, we also find that reserve prices are too clustered at zero and at multiples of €50,000 as to be consistent with fully rational behavior. If, as in our data, entitlement and clustering effects are persistent and quantitatively relevant, the option of choosing a reserve price might be an impediment to market efficiency as sellers set too high reserve prices resulting in too little trade. On the upside, our findings suggest that simple

[3]This paper is joint work with Florian Englmaier.

microeconomic theory gives us a lot of mileage in explaining market behavior in complex environments. We document that many sellers very finely adjust their behavior to demand patterns and try to strategically exploit potential arbitrage possibilities. Moreover, we are able to show that the adoption of heuristic pricing rules by the sellers does not affect the expected revenue from an auction dramatically, as long as the chosen reserve price lies below the optimal level and competition among the bidders is sufficiently intense.

The fourth chapter is based on the paper The Evaluation of Complex Goods - Evidence From Online Car Sales.[4] It adds to the results presented in Chapter 2 by analyzing whether individuals efficiently aggregate all relevant information on the constituent characteristics of a complex good in a situation, where their decisions involve large monetary stakes. In particular, we focus on the market for used cars, in which the basic situation is comparable to that of the buyers in the virtual Hattrick-economy: People are presented with many details of a complex good and have to form their valuation for it. Based on detailed field data on more than 80,000 used car offers in a large online vehicle marketplace, we find evidence for biased information processing also in this market. While the precise date of first registration is publicly and prominently stated for each car, we identify an amplified value adjustment for otherwise identical cars where the year-count changes. These distinct discontinuities in the price pattern indicate that individuals over-react to the figure displayed in the latter, while underrating the finer information on a car's age as conveyed through the month of first registration. While similar inefficiencies in the utilization of information have been substantiated for small stake purchases (Lee and Malmendier, 2007) and in financial markets (Gilbert et al., 2008), we are able to document that analogous inattentiveness also persists in situations involving large stake real money decisions, indicating that inattentiveness within markets can have sizeable economic consequences. More generally, it stands to reason that such effects also exist for other markets, or to use the words of Akerlof (1970): "*The automobile market is used as a finger exercise to illustrate and develop these thoughts. It should be emphasized that this market is chosen for its concreteness and ease in understanding rather than for its importance or realism.*" (cf. p.489)

[4]This paper is joint work with Florian Englmaier.

The last chapter is based on the paper Strategic Seller Actions in Auctions with Asymmetric Bidders. In this paper, we examine theoretically a setting where the auctioneer faces two asymmetric bidders. The novel feature of this model is that, in addition to well-established interferences into the competitive process like handicaps and affirmative action, it allows for another dimension of strategic choice for the auctioneer to support specific bidders and thereby alter the auction outcome to her benefit. In particular, Cantillon (2008) shows in a general framework that a reduction in the degree of asymmetry among the bidders increases the expected revenue for the auctioneer. Intuitively, asymmetry hurts the seller, as competition among bidders is reduced resulting in a lower expected final price. We use this finding to analyze its implications for the strategic scope an auctioneer may have once the auction format has been set and committed to. Employing a simple two-bidder second-price-auction setting, we examine the possibilities of the seller to manipulate the valuations of the bidders without violating the specified rules of the implemented auction, and analyze how she optimally acts to maximize her expected revenue. We find that favoring the potential "losing bidder" is an optimal strategy for the seller. Intuitively, by taking actions to support the weaker of two participating bidders, the seller can make them more competing rivals, which in turn leads to an increase of her expected revenue. Furthermore, we show that this result holds true even if this favoritism causes a negative impact on the valuation of the competing "strong" bidder. However, this results in ex post inefficiency in terms of social surplus, whenever strong bidder still wins the auction and the seller "invested" in the ex-post losing bidder. At a policy level, these results suggest that in merger analysis and public tenders involving auctions, interference with the competitive outcomes may occur through more subtle and indirect channels, other than handicaps or affirmative action, and should be accounted for. We also discuss some possible applications, where such opportunities may indeed arise in practice.

Chapter 2

Does Bidding for Complex Goods Reflect All Relevant Information? Field Evidence From Online Gaming

2.1 Introduction

In recent years auctions on the internet experienced a vast increase in popularity. Today, all kinds of products are traded on market-platforms like *eBay.com* or *amazon.com*. In contrast to retail markets, where the price is typically fixed by the selling party, the outcome of an auction directly stems from buyer-level evaluations. Many of the featured items are of complex nature to the extent that they exhibit multiple dimensions of quality, like computers, mobile phones, or even cars. We are especially interested in the behavior of individuals when they have to decide on how much to bid for such items: Do they incorporate all of the available information, or do they neglect important aspects in their evaluation of the good? Economic theory suggests that the final price fully reflects all pieces of available information that are pertinent to the item's value. In this paper we take a novel approach to empirically examine individual bidding behavior in a highly controlled auction environment, where all crucial information on the offered products is openly accessible. We provide strong evidence that bidders neglect substantial parts of valuable information, even if it is readily provided. We find that people systematically fail to efficiently aggregate the available information on specific attributes of the items on sale.

Our evidence is based on detailed field data from Hattrick (*HT*), one of the largest and most popular online games. We argue that the elaborate framework of this game provides an excellent data source to empirically address our main question. Founded in 1997, *HT* is a browser-based free online football manager game with over 950,000 registered users and is available in over 124 countries, mostly in the native language. The basic concept of the game is to manage your own virtual football club consisting of virtual football players. The tasks for the human managers are manifold combining the elements of economic management, tactical options, and community interaction. Alongside a sportive component, the competition with teams of other human participants, the game demands from the user to develop a sound financial strategy for his club.[1] Typically this includes to profitably train virtual players and to trade them with other managers on the game's internal transfer market.[2]

Since the virtual football players in the game resemble complex goods, *HT*'s transfer market exactly provides the very situation that is of interest to us: The managers have to decide on how much they are willing to pay for a specific player, while they are provided with very detailed information on his attributes. More precisely, they learn his full attribute vector and thus there is no information asymmetry between buyers and sellers. On this basis, each bidder has to independently estimate the market value, where his individual valuation can be affected by various other factors like a budget constraint or the acuteness to acquire a new player. Hence, the private value paradigm applies best to this situation. Conveniently, the selling mechanism implemented in *HT*'s transfer market resembles an English ascending auction and thus the winning bid reflects an individual buyer's evaluation. Our large sample of transactions from this market allows us to empirically analyze to what extent the various attribute properties of the virtual players account for their market values. In turn, we test if the provided information is reflected in the actual prices, or whether important parts of it were excluded from individual managers' bidding considerations.

We obtained detailed information on 17,510 virtual players aged between 17 and 20 years, all listed as keepers on *HT*'s transfer market between May 01, 2008 and May 15, 2008. A key feature to our identification strategy is that the players in our sample can be regarded

[1]Source: http://www.hattrick.org

[2]Henceforth we will refer to the human users as "managers", while using the term "player" to address virtual football players.

as close substitutes in the dimension of their skill levels.[3] For reasons, which we will lay out in more detail below, this allows us to identify the impact of the second-most influential element to a player's value, namely his age, which is represented in the form "X years and Y days". By the design of the underlying game algorithm, the effect from training ceteris paribus declines with age. However, a player who is just a few days older than another - being otherwise identical - should not be worth much more, since the difference in their total age, and thus in their potential training-benefits, is marginal. Moreover, the above reasoning remains true even if the one player already turned a year older while the other's birthday lies just ahead - a fact that was fully confirmed to us by the makers of *HT*. Our leading hypothesis is thus that the transfer price should decrease continuously with the total age as measured in units of days, an information that is readily available and explicitly stated to every potential bidder.

Our data tell a different story. In contrast to our predictions, the value of a player does not decline smoothly in his total age, but exhibits systematic discontinuities in the price pattern, a finding which we label the "birthday effect": Depending on whether players are sold one week before or after their birthday, we find highly significant differences of sizeable magnitude in prices paid. For example, ceteris paribus a keeper who just passed his eighteenth birthday loses up to 21% in market value compared to a seventeen year old, whose birthday lies just a few days ahead. However, the only real difference between them is a unit increase in the former's year-count. Importantly, and confirmed by the makers of *HT*, the behavior we observe cannot be rationally justified through anything in the underlying game algorithm. Our intuition is that even though the exact age is clearly stated in the form "X years and Y days" when they submit their bids, the managers largely focus on the years of age while insufficiently incorporating the finer information given in the days of age.[4] Though we incorporate the finer information of the precise age in day units in our estimation analysis, we find that the age-group still has a large and highly significant impact on the observed prices. However, the latter contains less and noisier information than the former and hence should not play any role at all. Thus, we find that redundant information

[3]Throughout the paper, we distinguish between a player's "skills" and "characteristics" in the set of his attributes. Pre-drawing on section 2.2.3, the term "skill" captures eight abilities that determine a player's type. To all other entries in the attribute vector we will refer to as "characteristics" or "traits".

[4]Figure 5 in Section 2.2.3 below shows an example for the typical in-game interface where a manager submits his bid for a player. The exact age is clearly stated among the first details given on the player.

is evidently influential, reversely implying that the managers in reality do not sufficiently incorporate the finer information on the precise age at hand in their decision making process. This establishes our main result: A majority of managers in *HT* obviously overreacts to the number displayed in "years" while at the same time underweights or simply ignores the finer information on this attribute as given by "days" of age. As a consequence, we observe substantial discontinuities and overly high bids for players that are about to turn older in the near future. This result remains robust across various specifications and additional controls.

In the second part of the paper, we identify possible driving forces behind our findings. Two possible explanations suggest themselves: First, if individuals are constrained in time, it might simply be too costly to search and cross-compare the prices of several players equivalent in terms of age. Second, the systematic bias in the bidding pattern we observe may be the outcome of the managers simplifying their decision making by generically adopting heuristic pricing rules. A small change in the search engine used to screen the transfer market of *HT* during September 2008 constitutes a natural experiment that allows us to distinguish between these possible explanations. While initially only the age-group was depicted in the post-search results overview, in the revised design now the precise age of a player is displayed. Since this reduces the time and number of clicks necessary to compare different players, it is reasonable to argue that this amounts to an exogenous reduction in the search costs of a buyer. Hence, if it is classic search costs that drives our result, if anything, the "birthday effect" should be mitigated in the revised design. We indeed find evidence that the discontinuities are attenuated relative to the situation before the change, but the price discontinuities do not fully disappear. We thus conclude that the effect is affected by but not solely explained through search costs. From the perspective of optimal market design, our results imply that careful considerations should be given to the way in which information is presented to the bidders, if even the cost for a few "clicks" to view some internet page suffices to trigger large discontinuities in the price pattern.

Our findings contribute to the analysis of how information affects economic decisions. Similar to our results, Pope (2008) shows that people overly base their evaluation of hospital quality on reported ordinal rankings, while additionally a more precise measure in form of a continuous quality score is stated. In the context of auctions on the internet, the existing literature mainly focuses on the comparison of different auction formats under differing information regimes (see e.g. Lucking-Reiley, 1999; Bajari and Hortacsu, 2004). We add to this branch of research by analyzing whether available information is used efficiently at all.

Moreover, a key feature of our findings is that they stem from novel data created in a highly controlled environment that allows us to derive complementary insights to existing studies on information usage in bidding behavior. For instance, Lee and Malmendier (2007) provide evidence for overbidding at the online-auction platform eBay. They find that bidders anchor on an irrelevant retail-store price of a good (if stated by the seller), while at the same time many of them neglect a lower price offered in form of a "buy-it-now"-option. In this paper, we demonstrate a similar information neglect on part of the managers in *HT* with respect to the age attribute. However, our data differs from theirs in several ways. First, in our framework all items on sale are presented in an exogenously pre-determined standard. Unlike in eBay, we therefore can rule out the mode of presentation and the informational content as an instrument for the sellers to take influence on buyer behavior. Second, there exists no outside market for the virtual players traded in *HT* that could potentially affect the prices on the transfer market. Third, the particular piece of information we find to be systematically neglected is explicitly presented to all bidders with certainty, while information like the "buy-it-now"-option in eBay are generally less easily accessible. Finally, the transfer market in *HT* exhibits no risk of post-auction default and is free of transaction costs. By design, we are thus able to rule out some of the common explanations brought forward in the context of biased bidding behavior. For instance, a non-standard utility of winning an auction, also referred to as "bidding fever", cannot explain our findings, since otherwise we would expect to observe overpayment for any arbitrary player in our sample, not just and systematically for those close to their birthdays.

We are well aware that there may be reservations to working with data from a source like *HT*, which is clearly labeled as a game and where all financial transactions are carried out in terms of virtual money. However, success in the game requires patience and a long-term planning horizon and according to the developers, an individual manager typically keeps playing actively for about three years. Moreover, they also state that as many as 500,000 managers visit their account every single day. Above that, even though the basic access to the game is free of charge, roughly 20% of registered managers not only devote much of their time to the game, but they also voluntarily invest real money in *HT* by opting for a costly premium account.[5] These stylized facts underpin our view, partially inferred from own introspection, that the *HT* managers are very ambitious and that the game provides rather strong incentives. Plausibly, the within game motivation can be regarded as high as that of participants in laboratory experiments or small stake internet sales. Moreover, in a recent study Castronova (2008) provides suggestive evidence that economic constraints like the law of demand apply similarly in virtual environments with virtual money as they do in the real world. It thus seems reasonable to argue that it is possible to retrieve meaningful insights on individual behavior also for real life situations involving real money.

The remainder of this paper is structured as follows. Section 2.2 provides a brief description of some basic concepts of *HT* and details on the transfer market, the virtual players, and the sample selection. The descriptive statistics of our data are presented in Section 2.3. In Section 2.4 the estimation model is specified and the main results from the empirical analysis are established. Section 2.5 discusses possible explanations and analyzes how a change in the game design affects our results. Section 2.6 concludes.

2.2 Institutional Background and Sample Selection

Our analysis is based on data on virtual players sold in Hattrick's transfer market. After a brief introduction into the game's basic principles, we discuss the implementation of its transfer market and bidding system.[6] We then provide more details on the goods traded, i.e. the virtual players, before we turn to a detailed description of our dataset.

2.2.1 Background on the Game and its Mechanics

Over the last few years games on the internet have become increasingly popular. More than ten years after its invention in 1997, *HT* is still among the most rapidly growing browser-based online games. The basic concept of the game is to manage your own virtual football club, combining the elements of economic management, tactical options, and community

[5] "Supportership" enables a package of further features and tools like bookmarks and statistics at a fee of about $30 per year on a non-subscription basis. The managers take no technical advantage in their in-game performance from this supporter status, but it merely "*make[s] your time at* Hattrick *easier as well as more fun*", as the operators of the game put it. (Source: http://www.hattrick.org/Help/Supporter/)

[6] All descriptions of the game relate to the time of our data collection.

interaction. To consistently perform well, it is necessary for a manager to utilize all three of these elements. Your team plays at least one weekly game in a national league system against teams coached by other managers. *HT* is played in semi-real time: While a match takes ninety real-time minutes to complete, a season in *HT* lasts for an in-game year that is normalized to 112 real-time days. The outcome of matches is determined by random simulation on the basis of the chosen strategies of the opponents, skills of the virtual players and other factors that determine the probabilities to win. The tasks for a manager to lead his team to success are numerous, ranging from decisions on match tactics and line-ups, over hiring team staff like doctors and co-trainers, over "drafting" a new player from the team's youth squad and either selling, keeping, or firing him as needed, to monitoring the team's training program. Many managers complete all of these tasks almost on a daily basis.

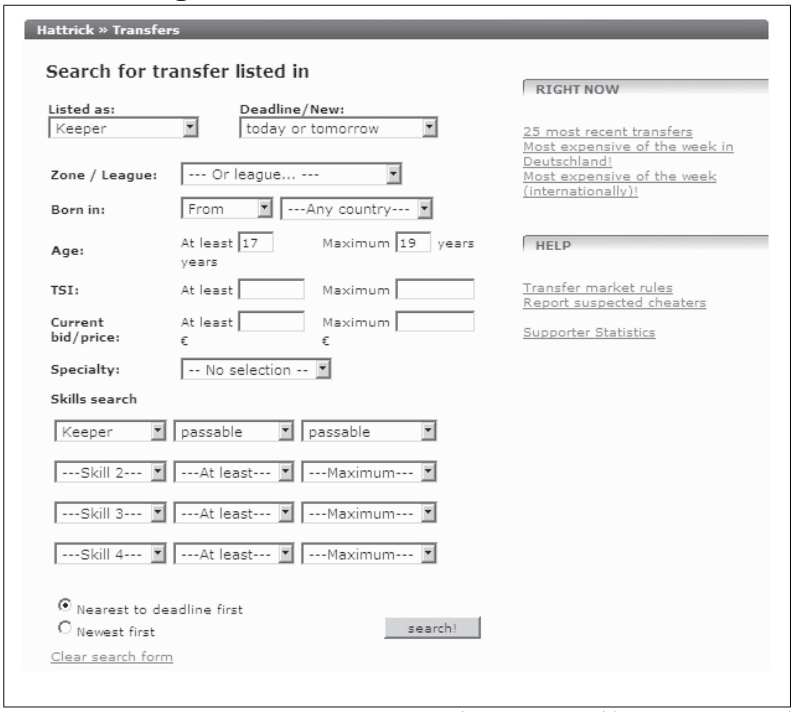
In addition to the sportive component, the game requires managers to develop a sound financial scheme for their clubs. The most promising way to raise (virtual) money in *HT* is to train and improve one's players and profitably sell them to other managers. Hence, trading players on the transfer market is a crucial element of the game. Financial success typically starts with specialization in a certain training strategy since the skills improve very slowly and only one out of eight can be trained at a time. For any skill, it takes several weeks until a "trainee" player "pops up" to the next higher level. Moreover, whether a player receives training at all and to what extent depends on the position he was lined-up in the last match he played. Although a complete description of the complex rules of the game even regarding only the training would be beyond the scope of this paper, the following example should illustrate the underlying mechanisms: Assume a manager chooses to train the skill *keeping*. In this case, only those players having played as goalkeepers during the week's matches will receive *keeping* training. Thus, any efficient training strategy requires considerable long-term persistence.

Usually, the managers stick to one particular training type and continually "produce" a specific player type. This is where the second key component to financial success comes into play, namely to profitable trade developed "trainees" and newly "drafted" youth players with other managers on the transfer-market. This latter task is of primary interest to us. It allows us to analyze how a manager evaluates a virtual player, essentially a multi-dimensional attribute vector, which in addition to his skills also contains other characteristics like personality traits, wage, and current form. Figure 2.3 in Section 2.2.3 presents a graphical illustration of a typical player profile in *HT* and the various attributes are discussed in more detail. For now, it is important to understand, that the quality of a player and his suitability to play certain positions in the team is completely determined by his attribute vector. Thus, the same is true for his market value. When offered on the transfer market, a player's complete attribute vector is freely accessible to any potential bidder, such that there is no information asymmetry between buyers and sellers. Importantly, for our analysis, on the web interface where the managers submit their bids, the age of a player is explicitly stated in the form "X years and Y days old" (see Figure 2.3).

2.2.2 Transactions: The Transfer Market

With an average of about 40,000 players offered for sale each single day, the transfer market in *HT* has a remarkable trading volume.[7] At any time, a manager can decide to offer one of his own players for sale. To do so, on the profile page of a particular player the manager sets a reserve price and hits a sell-button. After the submission of the sell order, the player is offered for exactly 72 hours from then on.

Figure 2.1: Transfer Market Search Mask



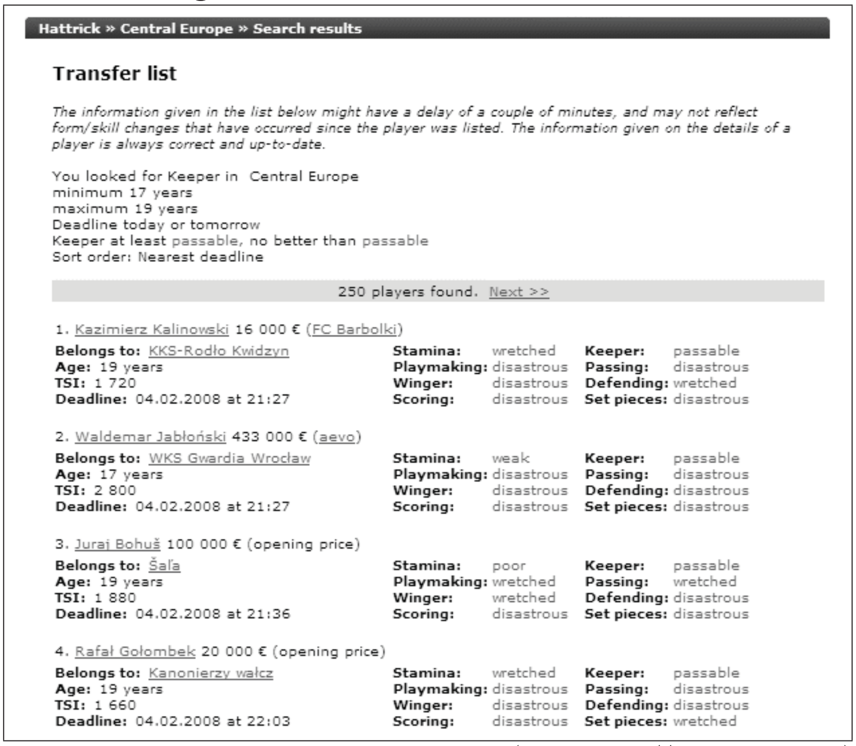
(Source: http://www.hattrick.org)

[7]Source: http://www.databased.at/HT/htpe

The selling mechanism implemented is an automatically extended English ascending open-bid auction: If someone places an offer less than 3 minutes before the deadline, the deadline will be extended for another 3 minutes. This continues until all bidders but one retire.

All players are displayed following an exogenous standard and the sellers have no means to affect the way how their player is presented to potential buyers. Figure 2.1 shows the typical user interface a manager is presented with when he enters the transfer market. It displays the search engine, which allows to filter for various player attributes like age, current bid, and - most importantly - up to four playing skills and their desired levels of ability.[8]

Figure 2.2: Transfer Market Search Results


Hattrick » Central Europe » Search results

Transfer list

The information given in the list below might have a delay of a couple of minutes, and may not reflect form/skill changes that have occurred since the player was listed. The information given on the details of a player is always correct and up-to-date.

You looked for Keeper in Central Europe
minimum 17 years
maximum 19 years
Deadline today or tomorrow
Keeper at least passable, no better than passable
Sort order: Nearest deadline

250 players found. Next >>

1. Kazimierz Kalinowski 16 000 € (FC Barbolki)
Belongs to: KKS-Rodło Kwidzyn **Stamina:** wretched **Keeper:** passable
Age: 19 years **Playmaking:** disastrous **Passing:** disastrous
TSI: 1 720 **Winger:** disastrous **Defending:** wretched
Deadline: 04.02.2008 at 21:27 **Scoring:** disastrous **Set pieces:** disastrous

2. Waldemar Jabłoński 433 000 € (aevo)
Belongs to: WKS Gwardia Wrocław **Stamina:** weak **Keeper:** passable
Age: 17 years **Playmaking:** disastrous **Passing:** disastrous
TSI: 2 800 **Winger:** disastrous **Defending:** disastrous
Deadline: 04.02.2008 at 21:27 **Scoring:** disastrous **Set pieces:** disastrous

3. Juraj Bohuš 100 000 € (opening price)
Belongs to: Šaľa **Stamina:** poor **Keeper:** passable
Age: 19 years **Playmaking:** wretched **Passing:** wretched
TSI: 1 880 **Winger:** wretched **Defending:** disastrous
Deadline: 04.02.2008 at 21:36 **Scoring:** disastrous **Set pieces:** disastrous

4. Rafał Gołombek 20 000 € (opening price)
Belongs to: Kanonierzy wałcz **Stamina:** wretched **Keeper:** passable
Age: 19 years **Playmaking:** disastrous **Passing:** disastrous
TSI: 1 660 **Winger:** disastrous **Defending:** disastrous
Deadline: 04.02.2008 at 22:03 **Scoring:** disastrous **Set pieces:** wretched

(Source: http://www.hattrick.org)

A search returns up to 250 offered players matching the selected filter, where an abstract of their main characteristics is displayed as shown in Figure 2.2.[9] Per default, results are sorted by closeness to deadline. By clicking on a player's name, his individual profile page opens displaying the full vector of attributes (see Figure 2.3). This is also the interface from which an actual bid can be submitted. Placed bids are binding and irreversible. Once the auction

[8]In the terminology of the game, skill levels are denoted as adjectives. To simplify the notation, we use integer values to address them, e.g. "passable" corresponds to score 6 of 20. Table 2.10 in Appendix 2.7.2 shows the detailed ranking, which can also be found in the game's manual.

[9]Observe that this preview only contains information on a player's age-group but not his precise age. In Section 2.5 we discuss why and analyze how a change of this display may alter our findings.

ends, the player is automatically transferred to the winning manager's team and the seller receives the winning bid.[10] Thus, in this framework there is no risk of post-auction default.

Having described the market from which our data stems, we now turn to a more detailed discussion of the goods traded in it, namely the virtual players and their attribute vectors.

2.2.3 Goods: The Virtual Players

The attribute vector of an individual virtual player has about thirty dimensions. Figure 2.3 shows the typical profile of a virtual player that is listed on the transfer market. Importantly, note that this is also the page where any prospective buyer submits his bid: To the right of the attribute vector the auction details are displayed and bids can be placed via the "bid!"-Button.

Figure 2.3: Virtual Player Profile

(Source: http://www.hattrick.org)

As we discuss in detail below, our leading hypothesis is that the value of a player should ceteris paribus continuously decrease with his total age as measured in day units. Therefore, we are especially interested in how this attribute is presented to the managers. As Figure 2.3 shows, literally topmost all information on the exact age is stated in the form *years* and *days*.[11] In addition, also the next birthday of the player is displayed, stating the actual calendar date when the player turns one year older. Factually, this attribute repeats

[10]Implemented to discourage excessive day-trading a small percentage of the price is deducted as a fee, which decreases with the time a player was member of a team.

[11]In the following, we use italics to denote the variable name in our data corresponding to an attribute.

the informational content contained in the variable *days* in yet another form. The precise information on the age attribute is thus clearly visible to any potential buyer.

Of major importance for the value of a player are the eight attributes displayed in the middle of Figure 2.3, which we denote as his "skills". While *stamina* and *set-pieces* are general skills, the remaining six - *playmaking, winger, scoring, keeping, passing,* and *defending* - determine a player's suitability to play in certain positions in the line-up. For instance, a player with his best skill being *keeping* is rationally classified as goalie. The player-skills are the only attributes that can be actively influenced by the manager via training and they drive the players' value to the largest extent. While in general the skills of a player are private information to his owner, if he is offered on the transfer market his full attribute vector becomes public information. Therefore, all attributes are freely accessible for any potential buyer. Table 2.1 provides an overview of the attributes and variables in our sample and displays the individual values they can take in the game.

Table 2.1: List of Variables

	Variable	Description	Range
Player attributes	years	Age in years (1 HT-year ≡ 112 real-time days)	17+
	days	Age in days (1 HT-day ≡ 1 real-time day)	{0,..,111}
	totalage	Total age in day units (contructed/normalized)	{0,..,448}
	days17-days20	Interaction term of days and age-group dummies	{0,..,111}
	form	Current form of player	{0,..,8}
	total skill index	Noisy indicator of overall quality of player	$\mathbb{N}_+$
	wage	Salary (exogenous; in virtual Euro)	$\mathbb{N}_+$
	keeper	Playing skill, position specific	{0,..,20}
	playmaking	Playing skill, position specific	{0,..,20}
	winger	Playing skill, position specific	{0,..,20}
	scoring	Playing skill, position specific	{0,..,20}
	passing	Playing skill, position specific	{0,..,20}
	defense	Playing skill, position specific	{0,..,20}
	setpieces	Playing skill for all player types	{0,..,20}
	stamina	Playing skill for all player types	{0,..,20}
	gentleness	High value if agreeable (ascending order)	{0,..,5}
	aggression	Low value if player aggressive (descending order)	{0,..,4}
	honesty	High value if honest (ascending order)	{0,..,5}
	player experience	Experience of player	{0,..,20}
	leadership	Leadership qualities of player	{0,..,7}
Auction Data	price	Auction end price paid by winning bidder	$\mathbb{N}_+$
	dtime	Time of deadline	hh.mm.ss
	dday	Day of deadline	dd.mm.yy
	buyer_countryID	Country of origin for buyer	$\mathbb{N}_+$
	buyer_searchcost	Proxy for buyer search cost based on broadband data	{0,1}
Dummy variables (1=yes, 0=no)	age17 - age20	Dummy for age-group	{0,1}
	peakhour	Auction deadline ended 5:30 p.m. - 10:00 p.m.	{0,1}
	mon - sun	On which day does the auction end?	{0,1}
	acquired	Proxy for previous sale (player countryID = seller countryID)	{0,1}
	expert	Proxy for buyer experience by leaguelevel	{0,1}

Among the remaining characteristics, a player's *total skill index* (*tsi*) is the one most likely to have a (positive) influence on market value, which represents a noisy measure for his overall abilities. To see this, note that *HT* calculates the skill-levels are as real numbers including hidden decimal places, the so-called "sub-skills", while the player profile only displays the adjective corresponding to the current integer value for each skill. With each training a player receives, the trained skill increases by a marginal increment (which is declining in age), and so does the *tsi*. While also correlated to other attributes (e.g. *form*), the *tsi* score thus constitutes a noisy signal for the sub-skills of a player, i.e. for how close he is to reach the next higher level in one of his skills.

A complete description of all characteristics is beyond the scope of this paper, but we employ the full set of attributes as control variables in our empirical analysis. For the players in our sample we have fairly clear predictions in which direction their effect on the price, if any, should go. All attributes not discussed are indeed of second-order importance to the value of the players in our sample and go into the direction which we would expect from the design and rules of the game.[12]

2.2.4 Sample Selection

A crucial feature of the game is that the players grow older over time. With increasing age, the marginal return from training declines and at around age thirty they slowly begin to deteriorate in their skills, such that they finally will have to be replaced. Therefore, each week a new cohort of players enters the game. Every week, each manager can "draft" one completely randomly created new player from a youth academy at a small fixed cost. The maximum level a skill can take for such a player is score 8 out of 20, but most commonly the highest skill will lie below that.[13] Like his whole attribute vector, also the player's age is randomly assigned and falls into the range between seventeen and twenty years. A player might for example be, say, "18 years 20 days" old when he is drawn.

[12] A test for joint significance of the control variables confirms that they have some impact, but the magnitudes of the individually significant coefficients are small.

[13] Score 6 for at least one of the skills is regarded as a minimum requirement by most managers to keep the player. For lower scores, the market value is close to zero and these players are usually instantly fired after they were drafted.

For our analysis, we concentrate on this specific subgroup of players for several reasons. First, a large majority of the newly drafted players has at most one skill at which they are reasonably good at and which determines their type. Second, young players comprehend what we label a large "advancement potential", since the effect of training on skill improvements is the larger, the younger a player is: All else equal a seventeen year old player will need fewer units of training to advance to the next skill level than if he was eighteen and so forth. In other words, the marginal training effect is the largest at the age of seventeen and is a decreasing function of age. As a consequence, only players up to the age of about twenty are regarded as appropriate "trainees" for a profitable training strategy and a way to raise money in the game. For higher ages, the market value converges to a lower bound which we label as a player's "consumption value" within the considered skill-region.[14] Finally, as a newly "drafted" player often does not fit into the chosen training strategy of his manager, these players are heavily traded on the transfer market.

By the nature of a football simulation, most of the several thousand players offered on the transfer market are of some field player type, like midfielder or forward. In *HT*, the performance of field players depends not only on one skill but rather on a combination of several different skills. Moreover, managers can assign individual tactical orders to field players and choose among various match tactics, both shifting the relative weights of specific playing skills, and hence their impact on overall performance.[15] As a consequence, the various skills of a player may receive very different weights in the individual evaluations across managers, making it hard to estimate their market value. To the contrary, goalkeepers are not affected by the chosen tactics, they cannot be assigned any individual orders, and *keeping* is by far the most important playing-skill. Hence, we naturally chose to focus on this subgroup of players for our analysis. In our identification strategy, this allows us to control for and back out the effect of the skill-component on the observed prices to the largest extent.

Summarizing, all players in the sample are aged between seventeen and twenty years and display an ability-level between score 6 to 8 out of 20 in the skill *keeping*, which accounts for the thickest market segment of players in this category. In addition, we restrict our sample to players without injuries, holding the health status constant.

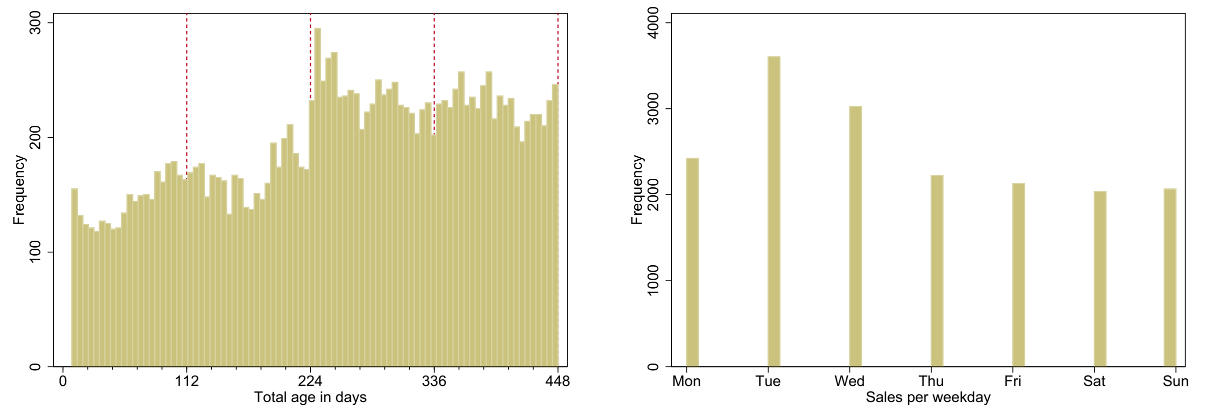
[14]See Section 2.4.1 for a detailed discussion of the market value of a player.

[15]For example, for a midfielder ordered to play "defensive" the *defense* skill becomes more important than if aligned "normal". Likewise, the tactic "pressing", for instance, increases the weight of the *passing*-skill for all field players, while it is not important for, say, defenders when the standard tactic is played.

2.3 Data Description

Our dataset consists of detailed information on 39,413 virtual goalkeepers that were traded on *HT*'s transfer market at any time of day within a consecutive collection period lasting from May 01, 2008 to May 15, 2008. In particular, we collected the full attribute vector for each single player along with details on the auction deadline, the final price, and the managers involved in the transaction.[16] For reasons laid out above, we restrict our focus to the subpopulation of players below age twenty-one, which yields 19,191 remaining observations. In addition, we exclude 1,510 players with a *keeping*-skill higher than score 8 and 171 outliers.[17] This leaves us with a final sample of 17,510 players.

Figure 2.4: Distributions of Age and Sales



(a) Distribution of total age (b) Distribution of sales per weekday

Figure 2.4a shows the age distribution for the players in our sample, indicating a roughly balanced distribution for *days* within each age-group. The distribution of sales per weekday is depicted in Figure 2.4b. On Tuesdays and Wednesdays we observe spikes in the number of sold players. Since new players can be "drafted" each Saturday after an weekly update and often are immediately offered for sale, the increased number of deadlines expiring on these days is not surprising. To examine whether *price* is affected by the auction end day, we add a dummy for each weekday as regressors on price.[18] We also control for the auction end time by including the dummy *peakhour* to indicate whether a player was sold between

[16]Refer to Table 2.1 for an detailed overview of all variables in our sample.

[17]Outliers are classified as prices above the 99%-percentile for each age-skill-combination. None of the results depends on their omission.

[18]For example, the effect of Saturday is measured by the dummy *sat*, which is 1 if the auction ended Saturdays, and 0 otherwise. Since it suffices to include six out of these dummies, we drop the dummy for Friday. Hence, the resulting coefficients are to be interpreted as relative differences with respect to Fridays.

5:30 p.m. and 10 p.m. , where the highest numbers of simultaneous online users are reached and most auctions expire.[19]

Table 2.2: Summary Statistics

Panel A. Overview

Variable	Obs.	Mean	Min.	Max.
price	17,510	159,393	1,000	1,500,000
years	17,510	19	17	20
days	17,510	55	0	111
totalage	17,510	245	3	447
total skill index	17,510	2,821	420	8,530
keeping	17,510	7	6	8
playmaking	17,510	1	1	4
scoring	17,510	1	1	5
passing	17,510	1	1	4
winger	17,510	1	1	5
defending	17,510	1	1	6
setpieces	17,510	2	1	11
stamina	17,510	5	1	9
leadership	17,510	4	1	7
wage	17,510	2,126	850	4,181
form	17,510	6	1	8
player experience	17,510	1	1	4

Panel B. Price for Age-Skill-Combinations

Skill	Age	Obs.	Mean	Min.	Max.
6	17	1,713	101,433	3,000	340,000
	18	1,956	45,469	3,001	125,000
	19	2,765	37,353	1,000	108,000
	20	2,661	34,699	2,000	100,000
7	17	1,296	384,871	99,500	1,129,005
	18	1,315	184,749	35,000	400,000
	19	2,006	150,489	60,000	350,000
	20	2,059	137,970	61,000	306,000
8	17	330	802,395	400,000	1,500,000
	18	474	572,709	225,500	1,062,000
	19	545	510,959	337,000	874,995
	20	390	487,770	235,000	800,001

Panel C. Prices per Age-Group and Skilllevel

	Variable	Value	Obs.	Percent	Mean	Min.	Max.
prices by age	years	17	3,339	19.07	280,724	3,000	1,500,000
		18	3,745	21.39	161,107	3,001	1,062,000
		19	5,316	30.36	128,599	1,000	874,995
		20	5,110	29.18	110,889	2,000	800,001
prices by skill-level	keeping	6	9,095	51.94	50,391	1,000	340,000
		7	6,676	38.13	198,876	35,000	1,129,005
		8	1,739	9.93	577,894	225,500	1,500,000

Notes: The variable $totalage \equiv 112 \cdot (years - 17) + days$ displays a player's precise age in day units. The minimum value of *totalage* at 3 reflects age "17 years and 3 days" and the maximum value at 447 equals "20 years and 111 days".

The summary statistics for our data are presented in Table 2.2. Panel A provides an overview of the most important player attributes, where the variable $totalage \equiv 112 \cdot (years - 17) + days$ is a normalized measure for the precise age of a player in day units, combining the information contained in the two variables *years* and *days*. Note that it has its minimum at value 3, since the youngest possible age a player can have is "17 years and 0 days" and each auction lasts for three days. With exception of *stamina* and *setpieces*, which are useful secondary skills for any type of player, the highest scores are attained in the *keeping*-skill, clearly classifying the players in our sample as "keepers". A correlation analysis of the main regressors on *price* yields a strong positive coefficient for *keeping* ($\rho = 0.80$) and *tsi* ($\rho = 0.79$).[20] Conversely, the age variables *years* ($\rho = -0.31$) and *totalage* ($\rho = -0.30$) have a significant negative correlation. Among the explanatory variables, by construction *years* and *totalage* evolve

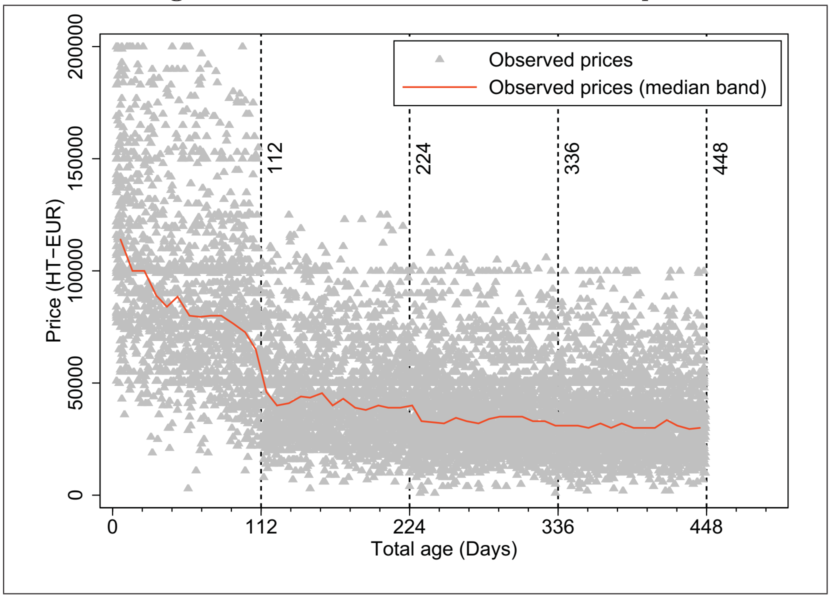
[19] See Figure 2.12 in Appendix 2.7.2.

[20] See Table 2.11 in Appendix 2.7.2. For the sake of clarity, only the most important attributes are depicted. Among the left-out attributes we find no unexpected correlations.

collinear at a degree of 0.96, while *days* has a correlation of 0.23. In addition, *keeper* and *tsi* co-move at $\rho = 0.87$. While in general collinearity among the explanatory variables can be problematic, our sample size is sufficiently large to produce precise parameter estimates.[21]

Returning to Table 2.2, listing the players by *years* in Panel C shows that the age-groups are roughly equally represented, with slightly less seventeen (19%) and eighteen (21%) year old players. The large majority of players has a *keeping* score of 6 (52%), while score 8 only accounts for 10% of the sample. In addition, Panel C also contains information on the price pattern for each age-group and skill level separately. Hardly surprising, the highest prices are paid for level-8 keepers and for the youngest players at the age of seventeen. Likewise, Panel B contains the price distributions for each combination of skill-level and age-group, yielding a first impression of how the price pattern evolves. For instance, a level-6 keeper aged seventeen on average yields €101,433 *HT*-currency, while the mean prices for age-groups eighteen (€45,469), nineteen (€37,353), and twenty (€34,699) are substantially lower. At a first glance, this sharp decline could be the result of a non-linear but continuous relation between *price* and *totalage*, which would be perfectly in line with our prediction that market price ceteris paribus decreases continuously in age.

Figure 2.5: Price Pattern for Level-6 Keepers



[21] Anticipating our results, all coefficients for the collinear regressors indeed turn out to be highly significant if jointly included in the estimation model.

However, a graphical inspection of the relation of price to the total age reveals intriguing patterns. Foreshadowing our main findings, the price pattern for level-6 keepers in Figure 2.5 displays strong discontinuities where the players turn one year older, a finding which methodologically neither a simple linear nor a non-linear model alone can explain. For level-7 and level-8 keepers qualitatively similar patterns arise,[22] implying that many managers systematically overpay for players on the verge to switch to the next higher age-group, obviously ignoring the imminent value loss he will experience on his birthday.

Having re-confirmed with the makers of the game that this “birthday effect”, as we refer to it, cannot be explained by anything in the underlying game algorithm, we take this finding as a first indication for inefficient usage of information on the precise age of a player as provided through the figure displayed in *days*.

2.4 Empirical Analysis

Given the richness of our data, we are able to analyze whether each player attribute is efficiently utilized in the managers' evaluations, i.e. whether it is correctly included or disregarded according to its predicted relation to price as implied by the rules of the game. By inspection, however, already the descriptive statistics suggest that this might not be the case for the attribute *days*, which conveys valuable information on a player's precise age. Hence, the identification of the relation between the different age attributes and price is at the core of our interest. Before we turn to an in-depth hedonic regression analysis, we briefly discuss the structural model our regression approach is based upon. In a further step we control for potential pitfalls in our data, discuss alternative specifications, and test for the robustness of our findings.

2.4.1 Estimation Model

In our identification strategy, we theoretically disaggregate market value into two distinct components. Above in Section 2.2.4 we already pointed out that the degree to which a player profits from training crucially depends on his age, a concept which we label as a player's

[22]See Figure 2.13 in Appendix 2.7.2.

"Advancement Potential Value" (APV). Recall that the younger a player is, the larger is his marginal benefit from training and the faster he improves in his skills. Thus, by design of the training algorithm, the APV should ceteris paribus steadily decline in the total age of a player. The second component, which we denote as the player's "Consumption Value" (CV), instead captures the extent to which he adds to his team's strength if lined-up for a match and is mainly driven by his skills. In combination, APV and CV account for a players market value:

$$Market\ Value = Consumption\ Value + Advancement\ Potential\ Value$$

A keepers' consumption value in HT is to the largest extent driven by his *keeping*-skill. While other skills and attributes like *stamina* or $form$ can be regarded as influential, importantly, his current performance - and thus his CV - is completely independent of age. Hence, a (simplified) specification for the CV of the players in our sample is given by[23]

$$CV = \alpha_1 + \beta_{keep} \cdot keeping + \beta_{stam} \cdot stamina + ... + \beta_{setp} \cdot setpieces + \beta_{form} \cdot form + ... + u_1 \quad (2.1)$$

In contrast, the advancement potential crucially depends on the age of a player. To put more structure on the estimation model for APV, review the observed price pattern in Figure 2.5. First, note that the observed discontinuities where the players turn one year older apparently differ in their relative size. To account for this possibility, we decompose the variable *years* into dummies for each age-group (*age17* - *age20*). For example, the effect of age-group eighteen is measured by *age18* taking value 1 and 0 otherwise. Since all four dummies are perfectly correlated, we need to include only three of them in our estimation model. Hence, by excluding the dummy *age17*, the resulting coefficients for the included dummies will capture the relative price differences for each age-group with respect to age seventeen. Second, the price pattern also shows that *within* each age-group the price-pattern declines gradually in *totalage*, just as we would expect. Note further, that the slope of this relation, i.e. the impact of a marginal day on price, remains roughly constant within an individual age-group. *Across* age-groups, however, the slope itself decreases. Formally, this corresponds to a piece-wise linear relationship between *totalage* and *price* being the true

[23]For the sake of clarity, we only include the most important attributes in this representation. In the regression analysis, the full vector of attributes was included.

underlying functional form in the population.[24] To control for possible changes in the slope, we interact the age-group dummies with *days*, thereby creating the variables *days17-days20* which display the age in days conditional on belonging to the specified age-group and zero otherwise. This allows us to identify the impact of a marginal day separately for each age-group. If the true underlying relationship was linear, the coefficients for the interaction terms would be the same over all age-groups. As a final step, recall that tsi is a noisy indicator of how close a player is to the next higher skill-level. Thus, this attribute is likely to be influential for the APV. Given these considerations, the estimation model for the APV can be described as follows:

$$\begin{aligned} APV &= \alpha_2 + \beta_{tsi} \cdot tsi + \beta_{age18} \cdot age18 + \beta_{age19} \cdot age19 + \beta_{age20} \cdot age20 \\ &\quad + \beta_{day17} \cdot days17 + \beta_{day18} \cdot days18 + \beta_{day19} \cdot days19 + \beta_{day20} \cdot days20 + u_2 \end{aligned} \quad (2.2)$$

In the empirical analysis we can only jointly estimate both value components, as reflected in the variable *price*. However, if we restrict our analysis to close substitute players with identical *keeping*-skill level, we are able to create partial homogeneity: By separating out the impact of the most influential attribute *keeping*, within each subgroup the consumption component of market value can be regarded as virtually constant. We are thus able to identify and estimate the APV component, which is of main interest to us, since it reflects the impact of age on price.[25]

Hypothesis 1 *If the market value of a player declines continuously in totalage, the value loss per year is fully captured through the aggregated marginal day-effects within this timespan. In the model framework, this is the case if and only if*

(i) the coefficient of age18 *equals the value decline per day in age-group seventeen* (β_{day17}) *times the number of days per year, i.e*

$$\beta_{age18} = 112 \cdot \beta_{day17},$$

[24]Having consulted with the makers of the game, this specification seems highly plausible and to be consistent with the ought-to devolution of prices. All our results remain qualitatively robust if we instead estimate a truly non-linear relationship. The regression results are available from the authors upon request.

[25]Since *keeping* is held constant, it cannot account for variation in *price*. Suppressing the influence of independent variables by holding them constant is a standard way to ensure statistical control. This technique allows us to infer whether an effect is due to one particular independent variable and not to another.

(ii) *the difference between the coefficients of* age19 *and* age18 *equals the value decline per day in age-group eighteen* (β_{day18}) *times the number of days per year, i.e*

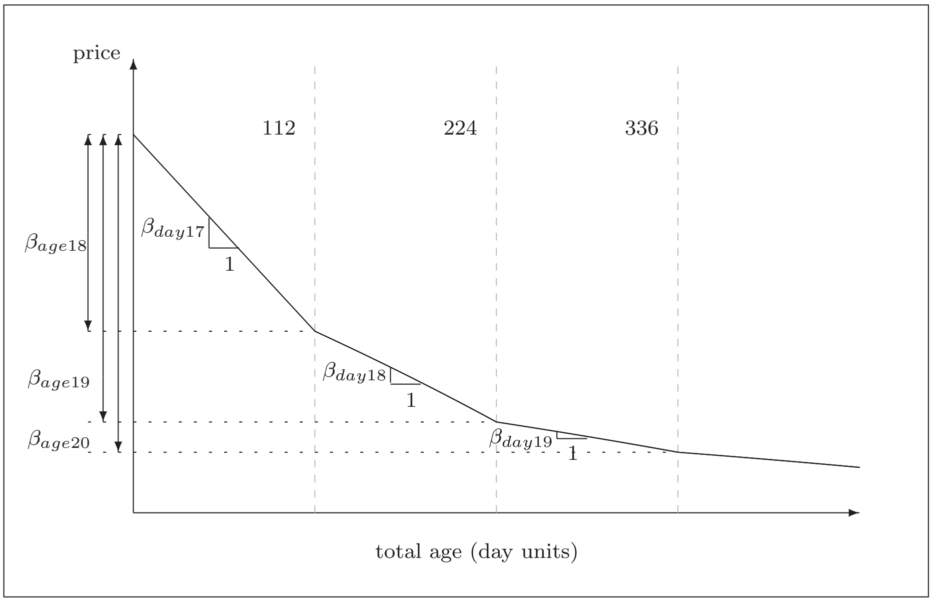
$$\beta_{age19} - \beta_{age18} = 112 \cdot \beta_{day18}, \qquad \textit{and}$$

(iii) *the difference between the coefficients of* age20 *and* age19 *equals the value decline per day in age-group nineteen* (β_{day19}) *times the number of days per year, i.e*

$$\beta_{age20} - \beta_{age19} = 112 \cdot \beta_{day19}.$$

For instance, consider a player who just turned nineteen. If market value declines steadily in *totalage*, then ceteris paribus his value should be equal to that of a player aged "18 years 0 days" net of the continuous value loss an average player experiences over 112 days within the age-group eighteen, respectively. This is implied by the testable predictions regarding the relations between the coefficients for the days-interaction terms and the age-group dummies as stated in the above hypothesis. Figure 2.6 contains a graphical illustration of the estimation model and the predictions from the hypothesis.

Figure 2.6: Structural Model and its Predictions



Intuitively, once we account for the effects on price of the finer information contained in the *days* of age, the noisier information as conveyed through the age-group (i.e. *years*) should be redundant and have no further impact on price. This is the case if and only if the coefficients on the age-group dummies solely reflect the steady day-by-day decline in market value that a player experiences within the previous age-groups.

2.4.2 Multivariate Regression Results

To test our main hypothesis (H1) formally, we conduct a hedonic regression analysis with the observed price as the dependent variable. For each *keeping*-level, we start out with a standard OLS procedure including only the main attributes from our theoretical estimation model presented above. In a second step, we incorporate a broad range of controls from the set of possible regressors as shown in Table 2.1, including all remaining skills (except of *keeping*) and all other player characteristics (e.g. *wage*, *leadership*, *honesty*,...). In addition, we also control for market fluctuations with regard to the daytime and weekday of the auction deadline. The regression results are depicted in Table 2.3, where each column represents one specification.

Table 2.3: Determinants of Price (OLS)

	Keeper = 6		Keeper = 7		Keeper = 8	
	I	II	III	IV	V	VI
age18	-86,498.41***	-88,096.64***	-324,311.89***	-325,926.07***	-438,805.74***	-455,444.61***
	(2,684.04)	(2,575.52)	(8,568.17)	(8,155.27)	(32,160.39)	(30,830.867)
age19	-97,509.31***	-100,614.80***	-366,515.16***	-368,241.79***	-502,696.46***	-524267.98***
	(2,599.52)	(2,513.81)	(8,189.75)	(7,793.75)	(31,098.47)	(30,064.959)
age20	-98,471.32***	-102,967.59***	-381,842.46***	-383,271.15***	-548,159.48***	-560,486.34***
	(2,597.10)	(2,538.99)	(8,172.21)	(7,836.20)	(30,979.14)	(30008.344)
days17	-597.40***	-624.86***	-2,557.94***	-2,614.91***	-3,223.88***	-3,128.98***
	(34.38)	(33.11)	(106.56)	(101.89)	(371.078)	(362.42)
days18	-38.43***	-58.32***	-234.73***	-255.88***	-400.73***	-215.21
	(13.84)	(13.09)	(45.55)	(42.59)	(155.024)	(158.30)
days19	9.83	4.40	-72.31**	-22.16	-248.85**	12.92
	(10.49)	(10.01)	(29.64)	(29.62)	(113.118)	(120.29)
days20	-13.31	4.70	-33.17	-8.29	174.14	227.80*
	(10.24)	(9.68)	(26.97)	(25.62)	(112.273)	(118.84)
tsi	23.36***	26.00***	45.58***	74.92***	58.51***	81.54***
	(0.68)	(1.61)	(1.64)	(3.09)	(3.253)	(5.12)
Intercept	87,099.03***	56,549.64***	370,202.16***	308,490.39***	736,388.85***	554,600.70***
	(2,521.54)	(5,547.66)	(8,582.69)	(13,734.94)	(34,122.57)	(44,567.45)
skills	no	yes	no	yes	no	yes
character	no	yes	no	yes	no	yes
daytime	no	yes	no	yes	no	yes
weekday	no	yes	no	yes	no	yes
R^2	0.56	0.61	0.72	0.75	0.59	0.64
N	9,095	9,095	6,676	6,676	1,739	1,739
F	511.55	185.19	768.58	249.93	188.55	68.26

Notes: Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. "Skills" captures the playing abilities except of *keeping* (= constant for each subgroup). "Character" contains all other player attributes except *tsi*. "Daytime" and "weekday" indicate whether dummies for daytime and day of the week were included.

In any specification, our estimation model predicts about 60% of the variability in the data, indicating a considerable degree of explanatory power. We start the derivation of our results with focus on column I in Table 2.3, which states the resulting coefficients and standard errors for the 9,095 level-6 keepers in our sample. Observe first that *tsi* has a positive effect on price with 99%-significant t-statistics, just as we would expect. Second, the coefficients for the age-group dummies are of large magnitude and significant on the highest level, indicating that there is indeed a declining relationship between age and price. For example, the negative coefficient β_{age18} (β_{age19}) indicates that the market value for a player aged "18 years 0 days" ("19 years 0 days") is on average about €86,498 (€97,509) less than that for a player aged "17 years 0 days". While improving the model fit in terms of a higher R^2, both the magnitude and significance of the coefficients remain virtually unaltered if we include further control regressors from the players' attribute vector and the auction details (column II).

As pointed out above, if the market value declines continuously in age, this substantial reduction will solely reflect the aggregated effects of the value loss per day over the duration of one year. In line with this argument, observe that the coefficients for *days17* and *days18* indeed reveal a significant negative effect of a marginal day on price within the age-groups seventeen and eighteen. However, this is not the case for the age-groups nineteen and twenty, immediately implying that from age nineteen onward the finer information on the age attribute as provided through the variable *days* has no significant impact on the market value of level-6 keepers.[26] This contrasts sharply with Hypothesis 1 and points towards an insufficient utilization of the provided information.

To substantiate this finding, we now turn to a test of our above predictions. Observe that in column II the aggregate day-by-day effect on price within the age-group seventeen amounts to $112 \cdot \beta_{day17} = 112 \cdot (-625) = -70,000$, which accounts only for 79.4% of the total reduction in market value as indicated by the coefficient $\beta_{age18} = -88,097$ in the model. Irrespective of the steady decline, on the day of his eighteenth birthday a player thus additionally loses €18,097 in market value. A Wald-Test confirms that this slump in value is highly significant (p-value: 0.000). Counterfactual to a fully continuous decline, the price pattern exhibits a substantial discontinuity at this point. Similarly, the total reduction of market value between ages "18 years 0 days" and "19 years 0 days" is given by $\beta_{age19} - \beta_{age18} = -12,518$, while the steady day-by-day decline only amounts to $112 \cdot \beta_{day18} = -6,496$, or 52% of the former. The remaining 48% ($-6,022$) establish another statistically significant discontinuity on the nineteenth birthday of a level-6 keeper (p-value: 0.000). Finally, also the third prediction

[26]In principle the *APV*-component could be close to zero already at that age-level. If that was the case, the observed prices would merely reflect the *CV*-part of market value and we ought to observe approximately similar prices for both age-groups, since their consumption values are independent of age. Yet, the prices in age-group twenty are significantly lower than that for players aged nineteen.

is not fulfilled, since the impact of *days19* is insignificant and hence the difference $\beta_{age20} - \beta_{age19} = -2,353$ identifies a third discontinuity located at the point where a player turns twenty (p-value: 0.002). Thus, the regression analysis validates the apparent discontinuities observed in the price pattern.

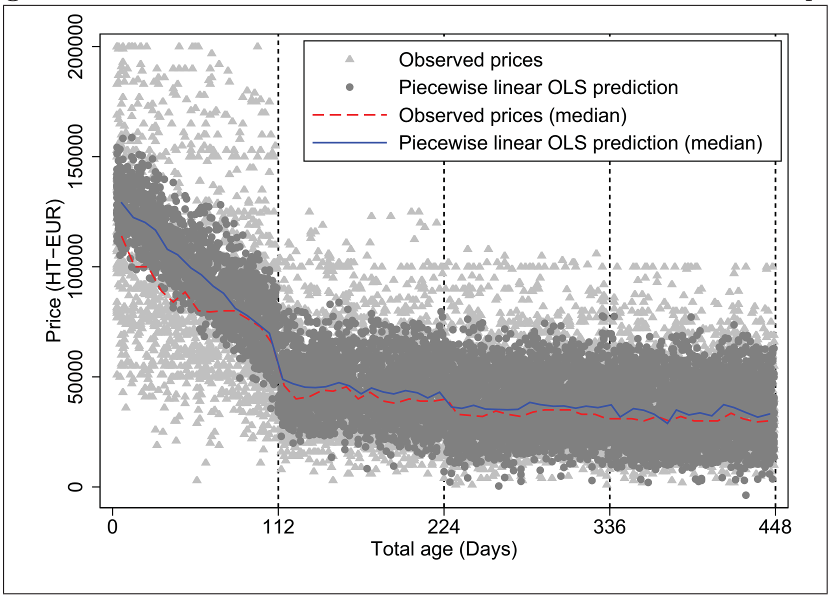
Table 2.4: The Birthday Effect - Size of Discontinuities

	Keeper = 6		Keeper = 7		Keeper = 8	
Birthday	I	II	III	IV	V	VI
18	-19634***	-18097***	-37816***	-33046***	-77718***	-104997***
	(0.23)	(0.21)	(0.12)	(0.10)	(0.18)	(0.23)
19	-6755***	-6022***	-15883***	-13644***	-18979	-68823***
	(0.61)	(0.48)	(0.38)	(0.32)	(0.30)	
20	-962**	-2353***	-7264***	-15029***	-17592*	-36218***
			(0.47)		(0.39)	

Notes: The table depicts the absolute magnitude of the discontinuities and their relative share of the total value decline in parentheses. If the day-by-day decline within an age-group was insignificant, the shown value reflects the total decline relative to the previous age-group (100%). Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

Moreover, the results we are able to demonstrate in our data prove to be highly robust if we instead analyze level-7 and level-8 keepers, either with or without controls. Table 2.4 provides an overview of the respective discontinuities for each regression approach from Table 2.3. Figure 2.7 illustrates the fit of our estimation model by plotting the predicted prices for level-6 keepers (with controls) against the observed transaction prices.

Figure 2.7: Model Fit - Predictions and Observations for Level-6 Keepers



Result 1 (Discontinuities in the Price Pattern): *The price pattern does not evolve continuously in the total age of a player but exhibits substantial and highly significant discontinuities at the players' birthdays. Hypothesis 1 can therefore be rejected.*

Once again, the total age of a player represents much finer information on his age attribute - and thus his advancement potential value - than it is the case for just the age-group. More precisely, after taking into account the exact information on the total age including the *days* - as we do in the model by including the interaction terms *days17*-*days20* - the variables *age18*-*age20* (and *years*, respectively) should be redundant for the pricing considerations of a manager as they bear no additional information. This contrasts sharply with the tremendous additional impact of the age-group dummies that we find in our data. With the moment a player turns one year older, his value slumps down dramatically. For some age-groups the loss on a single day, his birthday, accounts for more than half of the total loss during the period of one year. This implies that the managers systematically overpay for players close to their birthday. Corollary 1 thus establishes our first main result.

Corollary 1 (Inefficient Use of Information): *Managers in* HT *base their evaluations of virtual players on the noisier information contained in the variable* years *although they are provided with much finer information in the form of age in* days*. They do not or not efficiently incorporate important information in their pricing considerations.*

To illustrate the in-game economic impact of this finding, consider the following thought experiment. With 40,000 trades per day and about 950,000 registered users, the average manager buys about 5 players per season. Suppose an individual manager only buys seventeen year old level-6 keepers with their birthdays close by. He could save up to five times the value loss a level-6 keeper experiences on his eighteenth birthday, i.e. roughly €90,000, if he instead opted for players that just turned eighteen. With that amount he could afford to buy one additional average level-6 keeper from age-group seventeen (mean price €101,433), or even two additional players from age-group eighteen (mean price €45,469), respectively.

2.4.3 Robustness of Results

Having laid out our main result, we now provide a series of robustness tests. First, we present the results from alternative regression procedures addressing potential pitfalls in the data. Subsequently, we analyze whether the effects persist if we restrict to a subsample of expert managers. In addition, we relax our above assumption of a constant *keeping*-level by allowing for skill variations in a pooled regression approach.

2.4.3.1 Alternative Regression Procedures

In deriving our results we naturally control for potential pitfalls like multicollinearity and heteroscedasticity. As our sample size is sufficiently large, we do not find the first to be a problem. However, the large diffusion in the observed prices for seventeen year old players raises suspicions of having non-constant variance in the error terms.[27] In the above regressions we thus account for possible correlations of the residuals across observations by applying Huber-White-Sandwich-estimators to produce robust standard errors. An alternative remedy to heteroscedasticity is to perform a log-linear transformation of the dependent variable, e.g. using the natural logarithm of *price* instead of the raw values in the regression analysis. The resulting coefficients are depicted in Table 2.5.

Table 2.5: Determinants of Price - Log-linear OLS

	Keeper = 6		Keeper = 7		Keeper = 8	
	I	II	III	IV	V	VI
age18	-1.0415***	-1.0714***	-0.9530***	-0.9638***	-0.5703***	-0.5967***
	(0.0271)	(0.0254)	(0.0248)	(0.0193)	(0.0361)	(0.0344)
age19	-1.3103***	-1.3706***	-1.1925***	-1.2075***	-0.6805***	-0.7161***
	(0.0253)	(0.0240)	(0.0178)	(0.0168)	(0.0338)	(0.0326)
age20	-1.3510***	-1.4402***	-1.2853***	-1.2979***	-0.7678***	-0.7877***
	(0.0260)	(0.0252)	(0.0175)	(0.0165)	(0.0341)	(0.0331)
days17	-0.0054***	-0.0061***	-0.0060***	-0.0063***	-0.0038***	-0.0037***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0004)	(0.0004)
days18	-0.0008***	-0.0012***	-0.0013***	-0.0014***	-0.0007***	-0.0004
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
days19	0.0003	0.0002	-0.0004**	-0.0002	-0.0005**	0.0000
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
days20	-0.0002	0.0002	-0.0002	-0.0001	0.0003	0.0004*
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
tsi	0.0004***	0.0005***	0.0002***	0.0003***	0.0001***	0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Intercept	10.8397***	10.0712***	12.4384***	12.0982***	13.3418***	13.0222***
	(0.0295)	(0.0886)	(0.0257)	(0.0496)	(0.0400)	(0.0620)
skills	no	yes	no	yes	no	yes
character	no	yes	no	yes	no	yes
daytime	no	yes	no	yes	no	yes
weekday	no	yes	no	yes	no	yes
R^2	0.48	0.56	0 .70	0.74	0.58	0.63
N	9,095	9,095	6,676	6,676	1,739	1,739
F	1,229.02	411.34	1,983.55	661.60	279.75	99.44

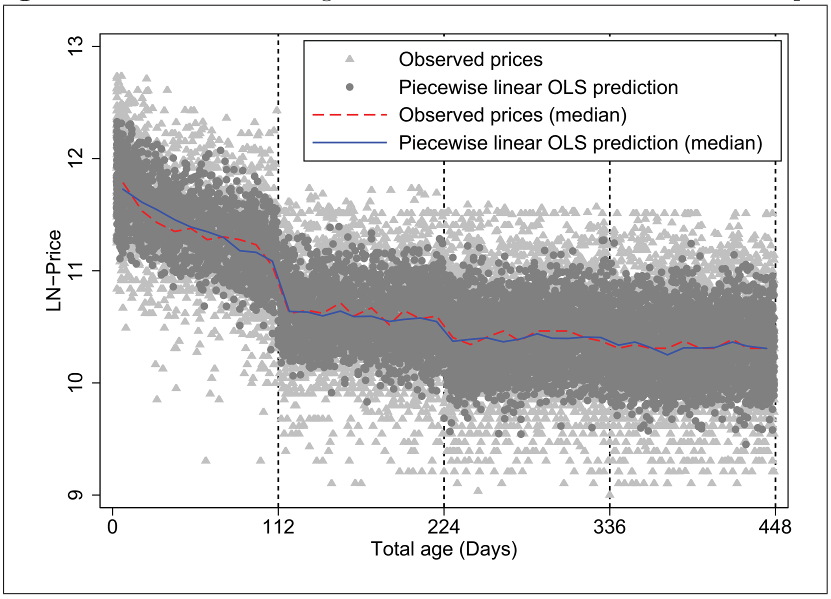
Notes: Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

The regression results qualitatively mirror those from Table 2.3. As before, the models yield reasonable levels of predictive power in terms of R^2, and across all regressions the quality indicator *tsi* has a significant positive impact on price. We representatively focus

[27]These worries are confirmed by both a Breusch-Pagan and a White test, both indicating heteroscedasticity in the data.

on the results for level-6 keepers presented in column II. All age-group dummies and the variables *days17* and *days18* have negative coefficients with 99%-significant t-statistics. A test of the predictions from Hypothesis 1 confirms the existence of the "birthday effect". To see this, note that the accumulated day-by-day decline in age-group seventeen is given by $112 \cdot \beta_{day17}$=-0.6832, explaining only about 64% of the total reduction in value between ages "18 years 0 days" and "17 years 0 days" as given by the coefficient of -1.0714 for *age18*. The difference validates a discontinuity on a player's eighteens birthday (p-value: 0.000), roughly accounting for a loss of 36% in market value.[28] Moreover, also the discontinuities upon entering age-groups nineteen and twenty prove to be significant at the 1%-level. Observe further that the fit of the predictions from the regression underlying column II with the true price pattern is very high, as graphically illustrated in Figure 2.8. An analysis for level-7 and level-8 keepers yields virtually identical results.

Figure 2.8: Model Fit - Log-linear OLS Predictions for Level-6 Keepers



To control for the possibility of influential outliers, we also perform a regression approach using an iteratively re-weighted least squares procedure for both absolute and log-transformed prices. By this method, each observation is assigned a weight $\omega \in [0, 1]$, where higher weights are given to better behaved observations and extremely deviant cases are excluded from the analysis. All results from the standard OLS regressions fully carry over to this approach and are shown in Tables 2.12 and 2.13 in Appendix 2.7.2.

[28] The exact percent difference is given by $100 \cdot (e^{\beta_{age18} - 112 \cdot \beta_{day17}} - 1) \approx -32\%$, where e is Euler's number and the exponent is the additional reduction in value that is not explained by the continuous day-by-day decline.

2.4.3.2 Experience

We have demonstrated that managers substantially overpay for players on the verge to enter the next higher age-group. Given the complexity of the game, one possible explanation for this finding could be linked to the experience of managers. It may well be that it is only the inexperienced users giving too less weight on the exact information on age, thereby overlooking the imminent birthday and the accompanying drop in market value.To test whether this is the case, we repeat our above analysis exclusively restricting to trades involving managers qualifying as “experts”. The latter are identified according to a proxy based on the division-level of each winning bidder’s team in our data.

Intuitively, every new manager starts out in the bottom division of a pyramid system, where each division is subdivided into leagues of eight competing managers. While the top division consists of only a single league, the second and third divisions already comprehend four and sixteen leagues, respectively. Depending on the number of registered users, there can be up to eleven divisions in a country, then with 4,096 leagues and up to 32,768 managers just in the bottom division. We thus classify a manager as an expert if his team plays in a division above a threshold-level that is defined such that on average about 20% of the managers in each country qualify as experts. For example, Germany is represented with ten divisions with a total capacity for 84,648 teams. To qualify as an expert, a German manager needs to play in division seven or higher, i.e. he must be among the best 19,112 or top 23% of all teams. Since starting in division ten, a manager needs at least three complete seasons to reach this threshold, which is a considerable time to gain experience.

The dummy *expert* takes value 1 if the winning bidder’s division met the respective expert-threshold in his country and value 0 otherwise. 3,538 auctions, or about 20% of all transactions in our data, involved experts according to this proxy.[29] Table 2.6 presents the respective regression results for this subgroup. First, note that both the magnitude and significance of most coefficients remains virtually unchanged compared to those from Table 2.3. All coefficients of the age-group dummies are significant at the highest level and resemble the ones from the full sample remarkably closely. Also the impact of the interaction terms *days17* and *days18* on price are of similar order and qualitatively go into the right direction, though

[29]Since our proxy on average classifies about 20% of the managers as experts, this share indicates a quite balanced representation of experienced managers in the analyzed segment of the transfer market.

β_{day18} is not always statistically significant, which is most likely due to less precise estimates obtained from a considerably lower number of observations.

Table 2.6: Determinants of Price - Expert Managers (OLS)

	Keeper = 6		Keeper = 7		Keeper = 8	
	I	II	III	IV	V	VI
age18	-85,446.95***	-88,888.14***	-325,942.80***	-336,094.47***	-486,899.09***	-497,561.66***
	(5,146.70)	(4,840.61)	(11,989.62)	(11,428.99)	(47,408.33)	(46,115.35)
age19	-96,192.23***	-99,711.65***	-371,658.35***	-379,088.67***	-547,177.09***	-573,735.24***
	(4,890.20)	(4,554.77)	(11,811.68)	(11,053.80)	(47,199.97)	(46,174.27)
age20	-92,262.77***	-100,641.45***	-379,658.00***	-388,504.79***	-598,629.31***	-613,268.10***
	(4,957.16)	(4,765.90)	(11,692.57)	(11,106.25)	(47,421.47)	(46,553.08)
days17	-526.99***	-563.51***	-2,595.46***	-2,694.19***	-3,725.46***	-3,652.63***
	(62.77)	(58.16)	(141.19)	(134.19)	(561.11)	(542.66)
days18	-45.74	-42.78	-293.14***	-260.51***	-348.24*	-193.17
	(44.30)	(43.23)	(86.16)	(81.42)	(202.15)	(233.48)
days19	32.57	33.05	-133.99	-7.20	-300.62	0.59
	(42.72)	(38.47)	(86.33)	(82.35)	(190.65)	(209.33)
days20	-43.27	21.45	-130.57*	-73.89	124.05	264.68
	(39.39)	(40.51)	(75.28)	(74.00)	(197.67)	(232.64)
tsi	37.16***	32.38***	66.66***	100.64***	56.59***	83.32***
	(2.23)	(5.20)	(4.04)	(6.90)	(5.43)	(8.02)
Intercept	54,447.79***	14,064.85	305,168.96***	250,519.79***	791,009.13***	677,286.89***
	(5,645.09)	(16,755.03)	(14,935.24)	(29,209.48)	(53,837.89)	(69,357.10)
skills	no	yes	no	yes	no	yes
character	no	yes	no	yes	no	yes
daytime	no	yes	no	yes	no	yes
weekday	no	yes	no	yes	no	yes
R^2	0.59	0.66	0.74	0.78	0.61	0.67
N	1,156	1,156	1,670	1,670	712	712
F	125.95	44.59	344.69	116.64	82.46	32.17

Notes: Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

Likewise, we find strong evidence for significant discontinuities in the price pattern if we test the predictions from Hypothesis 1. For instance, in column II the market value of an average level-6 keeper slumps down by $\beta_{age18} - 112 \cdot \beta_{day17} = -25,720$ on the day of his eighteenth birthday, which accounts for 29% of the total reduction, or for 41% of the aggregated day-by-day effect. Across all columns, a Wald-test confirms that the discontinuities between seventeen and eighteen are highly significant at p-values ≤ 0.00. Since *days18* has no significant effect on price in column II, to validate the existence of a discontinuity on the nineteenth birthday it would suffice to verify that the coefficients of *age18* and *age19* significantly differ from each other (which they do at a p-value of 0.002). If we despite the statistical insignificance account for the indicated day-by-day loss in age-group eighteen, i.e. $\beta_{day18} = -43$, this still yields a significant discontinuity of size $-6,008$ or 56% of the total decline (p-value: 0.090). For level-7 and level-8 keepers the results are qualitatively similar. Though we find no significant birthday effect for age-group twenty, which is also likely due to the reduced number of observations, the analysis clearly indicates that also expert managers fail to fully incorporate the information contained in the variable *days*. In any case, we can conclude that it is not lack of experience that drives our results.

Result 2 : *The birthday effect cannot be explained by a lack of experience.*

2.4.3.3 Skill Variations

So far we backed out the impact of the most influential skill by comparing players with a constant *keeping*-score. For a final robustness check we relax this restriction by pooling all observations and conduct regressions for the full sample. To control for the effect of skill variations on the market value of a player, we include the dummy variables *keeper7* and *keeper8* taking value 1 to indicate level-7 and level-8 keepers, respectively, and 0 otherwise. The regression approach yields the results shown in Table 2.7.

Table 2.7: Determinants of Price - Skill Variations (OLS)

	I	II
age18	-196,849.12***	-199,496.02***
	(5,733.18)	(5,599.87)
age19	-224,998.64***	-229,919.40***
	(5,573.30)	(5,447.43)
age20	-235,149.77***	-240,662.03***
	(5,579.50)	(5,483.80)
days17	-1,375.48***	-1,489.60***
	(71.53)	(68.92)
days18	-152.43***	-230.37***
	(27.00)	(26.42)
days19	-58.08***	-68.94***
	(19.48)	(19.31)
days20	-7.94	-2.90
	(18.91)	(18.14)
tsi	44.50***	72.02***
	(1.27)	(2.32)
keeper7	89,272.31***	64,041.62***
	(1,679.74)	(2,079.32)
keeper8	395,545.22***	320,797.67***
	(4,476.95)	(5,984.13)
Intercept	157,449.48***	183,130.47***
	(5,485.62)	(9,478.74)
skills	no	yes
character	no	yes
daytime	no	yes
weekday	no	yes
R^2	0.86	0.87
N	17,510	17,510
F	4,926.05	1,710.14

Notes: Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

Consistent with the theoretical considerations, the *keeping*-score has a strong positive and highly significant effect on the market value. Relative to an average level-6 keeper, the prices paid for level-7 and level-8 keepers are substantially higher. However, if we compare the coefficient for *keeper7* to that for *age18*, we find that the average price increase due to an skill improvement from level 6 to level 7 in *keeping* is relatively smaller than the average age-induced reduction between age "17 years 0 days" and age "18 years 0 days". In contrast, the coefficient on *keeper8* has by far the largest magnitude, which implies a non-linear price-pattern across skill-levels. Among the age-regressors, all coefficients are negative with 99%-significant t-statistics except for *days20*. This again confirms the importance of age for the market value of the players in our sample. Representatively focusing on column II, a test whether the price pattern evolves continuously further confirms the robustness of our previous results. We find significant downward jumps of additional 19% (p-value: 0.0000), 18% (p-value: 0.0243), and 39% (p-value: 0.0852) relative to the aggregated day-by-day decline at the eighteenth, nineteenth, and twentieth birthday of a player, respectively. Hence, even if we control for variation in the *keeping*-skill, the birthday effect is highly persistent.

For the sake of clarity, we only state the most important variables in the above regressions.[30] In general, the transfer market adheres to standard economic findings: A high supply of players on Tuesdays and Wednesdays leads to significantly lower prices, while on Saturdays and Sundays, where many managers are online simultaneously and imply a high demand, competition is fiercer among the buyers and thus the winning bids are somewhat higher. More importantly, however, it is worth pointing out that the average manager shows a rather sophisticated bidding behavior, indicating that they actually try to thoroughly elicit the value of a player they bid for. In no specification any attribute that, according to the rules of the game, is irrelevant to a keeper's market value had a statistically significant impact on *price*. For instance, among others a keeper's value is not affected by, say, his *playmaking* or *scoring* abilities, and consistent with this intuition all our regressions indicate that the managers correctly exclude these from their pricing considerations. Similarly, throughout all specifications we find a small but significant positive influence for the skills *stamina* and *defense*, which are of second-order importance for keepers.

[30] A detailed overview of the regression results is available from the authors upon request.

2.5 Possible Explanations for the Birthday Effect

In light of a rather sophisticated understanding of the game as displayed with respect to other attributes, the main finding we are able to document in our data seems even more puzzling. If bidders are careful enough to check out numerous details of the attribute vector of a player, why do they systematically pay too little attention to the valuable information conveyed through the age in days? Though the managers do not disregard the impact of precise age as indicated by the evidential continuous decline *within* the age-groups, they fail to recognize the connection to subsequent or previous age-groups. Our intuition is that the managers evaluate players relative to the average player from the same age-group, while the more relevant and informative peer group consists of players of close-by $totalage$-levels, irrespectively of the age-group the latter belong to.

To illustrate what we have in mind, consider a manager who has to evaluate a player with given attributes aged “17 years 105 days”, i.e. one week before his eighteenth birthday. All else equal, to elicit how much to bid for this player, he should look up and compare the prices for players of a similar total age, say, roughly from two weeks younger (“17 years 91 days”) to two weeks older (“18 years 7 days”). To get this information the manager has to screen a large number of players on the transfer market to find enough falling into that age-range. Importantly, recall from Figures 2.2 and 2.3 that a manager has to visit each offered player’s profile to learn his precise age. To find enough players in the relevant peer group thus involves a time consuming and thus costly search. This implies that one possible explanation for the birthday effect lies within the design of the user interface of the search engine, which is used by the managers to screen the market: It may be physical search costs that prevent a manager from efficient information aggregation. Before we address the impact of search costs empirically, we briefly outline a very simple theoretical approach relating costly search to the evaluation of the virtual players. In addition, we also discuss other possible factors that could explain the birthday effect.

2.5.1 Search Costs

2.5.1.1 A Simple Search Cost Dependent Evaluation Model

Consider a risk-neutral manager j who wants to evaluate a particular player $i = (k_i, y_i, d_i, X_i)$, where $k_i \in \{6, 7, 8\}$ denotes his keeping skill, $y_i \in \{17, 18, 19, 20\}$ his age-group, $d_i \in \{0, ..., 111\}$ his days of age, and X_i all other attributes of the player, respectively. Normalize by $a_i = 112 \cdot (y_i - 17) + d_i$ the total age in days. For given values of $k_i = \overline{k}$ and $X_i = \overline{X}$, the manager j's value estimate for player i in dependence of his age attribute is described by the function $E_j[v_i] : (y_i, d_i) \to \mathbb{R}^+$.[31] More specifically, let

$$E_j[v_i] := (1 - \pi(c_j)) \cdot \overline{v}_y^j + \pi(c_j) \cdot v_{a_i}^j,$$

where $\overline{v}_y^j$ is the value of an average keeper of age-group y to manager j, and $v_{a_i}^j$ denotes his precise value of player i. For simplicity, assume that $\overline{v}_y^j$ is commonly available free of cost. His value estimate is a convex combination of the average value and his true value, where the relative weight $\pi(c_j)$ is a function of his search costs c_j. By screening the transfer market for otherwise identical players within an age-range around a_i, he can learn their values and thus increase the weight $\pi(\cdot)$ on his true value for player i and thereby obtains a more precise estimate.[32] Generally, the intensity of this search will depend on how costly, or time consuming, it is to find appropriate players in the respective age interval. Formally, assume that the convex weighting function $\pi(c_j)$ has the following properties:

$$\begin{aligned} \lim_{c_j \to \infty} \pi(c_j) &= 0 \\ \lim_{c_j \to 0} \pi(c_j) &= 1 \\ \pi'(c_j) &< 0 \quad \forall \ c_j \in \mathbb{R}_0^+. \end{aligned}$$

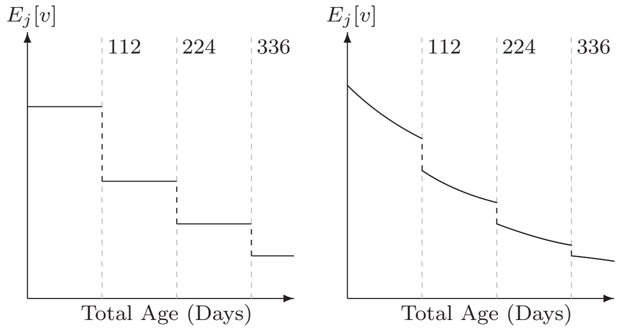
Note that in a second price auction he will bid exactly $b_j = E_j[v_i]$. We thus can distinguish three scenarios. First, consider that the search costs are sufficiently large such that $\pi(c_j) = 0$. Then manager j's bid will reflect the average value $\overline{v}_y^j$. Second, for a given c_j suppose that $\pi(c_j) < 1$. If $v_{a_i}^j > \overline{v}_y^j$, the manager bids too low and is less likely to win the auction, though his true value for the player would be higher than his estimate. Conversely, if $v_{a_i}^j < \overline{v}_y^j$, he will bid above his true valuation for player i. While the former case is unproblematic, in the

[31] To simplify the notation we suppress $\overline{k}$ and $\overline{X}$ in the expressions.

[32] The underlying rationale may be best explained by assuming that for any c_j, the manager solves an optimal search problem, which determines the number of players he optimally screens. In turn, this implicitly determines the extent to which he learns $v_{a_i}^j$.

latter the manager with the least precise estimate will determine the final price. Third, in the absence of search costs, manager j will fully learn his precise value, i.e. $\pi(0) = 1$. These cases are illustrated in Figure 2.9.

Figure 2.9: Expected Valuation in Dependence of Search Costs



(a) $\pi = 0$ **(b)** $0 < \pi < 1$ **(c)** $\pi = 1$

For lower search cost $c'_j < c_j$, the estimates of any individual manager j should become more accurate in the sense that they become closer to his precise value $v^j_{a_i}$ since $\pi(c_j) < \pi(c'_j)$. In the following we test this prediction empirically, to analyze whether the birthday effect can be explained by search costs.

2.5.1.2 Broadband Access as a Proxy for search costs

To get a first impression whether the birthday effect can indeed be attributed to search costs, as an initial coarse approach we construct a proxy for an individual buyer's search cost by matching the information on his country of origin with country-level data on high-speed internet subscriber rates per 100 inhabitants.[33] Intuitively, it is reasonable to assume that the search cost for obtaining the information on the exact age of a player, i.e. the cost of a single "click" to view the player's profile, are larger for managers with a slow internet connection because it takes longer to load each page. To see whether higher search costs exacerbate the birthday effect, we test for differences between the coefficients estimated over the group with high search costs and those estimated over the group with low search costs. Qualitatively consistent with our argument, for buyers from countries with mainly slow internet connections, or likewise high search costs, we find a tendency for an increased magnitude of the discontinuities. However, these differences are statistically insignificant which is not surprising given the coarse nature of our proxy.[34]

[33]The data is provided through the International Telecommunication Union (ITU) at http://www.itu.int.
[34]See Appendix 2.7.1 and Table 2.14 in Appendix 2.7.2 for the details of this approach.

2.5.1.3 Change of Game Design - A Natural Experiment

A subtle change in the design of *HT*'s transfer market enables us to more directly address the question whether search costs are at the core of our findings.[35] In particular, recall from the discussion of the transfer market in Section 2.2.2 that a search for desired players delivers an overview of suitable offers matching the selected search filter. This overview already contains a preview on the players' attributes, among others including the values for *tsi* and the full set of skills. Yet, as shown in Figure 2.2 above, during our initial data acquisition with respect to the age of the players only the age-group was displayed, i.e. the variable *years*. Thus, the preview conveyed only partial information on the age attribute and consequentially a manager was forced to inspect a player's full profile to also learn about the precise age, i.e. the information on the *days*.

During November 2008, however, *HT* implemented a general design makeover, which also included small, but for our purposes highly appealing changes of the transfer market. Particularly, as shown in Figure 2.10, in the revised design the search result overview now displays the precise age in the preview for each player, i.e. both variables *years* and *days*. Intuitively, this reduces the time and number of clicks necessary to compare different players in the appropriate peer group. Thus, it is reasonable to argue that this amounts to an exogenous reduction of the search costs for buyers.

In addition, also the search engine of the transfer market was slightly altered. In the old system, each search inquiry was restricted to one out of eight geographic zones at a time, but unrestricted with respect to filters on the attributes of the players. Within the new design, the zone restriction ceased to apply. While this change affects all ages equally and thus has no direct effect on our results, it implies a reduction in the number of separate search rounds necessary to screen the whole market supply for a particular type of player. At the same time, the operators introduced a limit on the number of age-groups that can be collectively searched at a time.[36] Applied to the players in our sample, the managers can now screen at most two consecutive age-groups together, i.e. players aged seventeen and eighteen, or eighteen and nineteen, or nineteen and twenty. While it is still possible to limit the search

[35]In fact, the alluded changes may have partly resulted from our extensive discussions with the makers of the game.

[36]According to the operators, this new restriction was implemented to prevent excessive server load from search inquiries to the transfer market database after removing the zone constraint.

Figure 2.10: Transfer Market Search Results - Revised Design

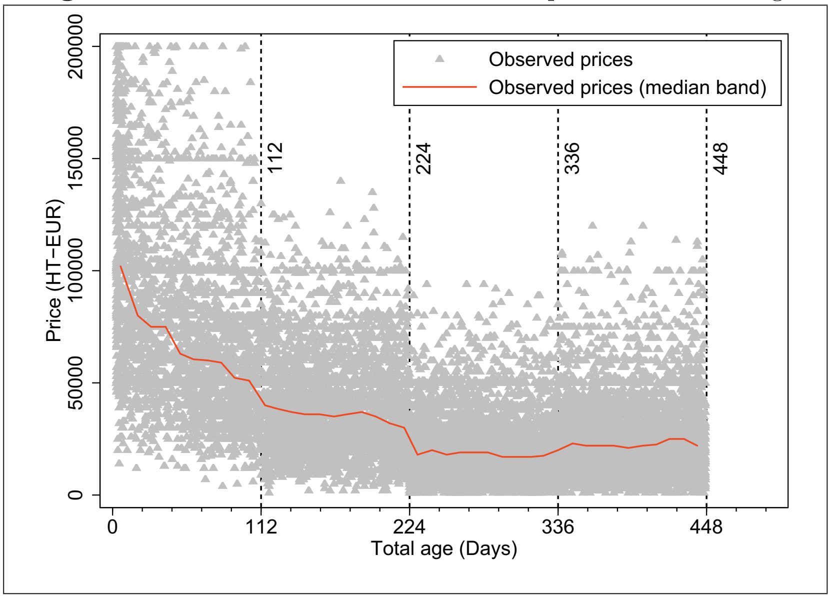
(Source: http://www.hattrick.org)

to one individual age-group at a time, the two-age-group setting is the preselected standard filter when a manager enters the transfer market. Our intuition is that this filter restriction potentially encourages a stronger focus on the relevant peer group, since it increases the likelihood that the search result list states players from the next higher age group right next to those from the lower one whose birthday is just imminent. Thus, if it is search costs that drives our result, if anything, the birthday effect should be mitigated by these changes.

With courtesy of the operators we obtained new data from the transfer market after the revised design was implemented. The sample includes detailed information on 30,295 keepers in the relevant skill- and age-groups that were sold during four consecutive weeks between December 11, 2008 and January 10, 2009.[37] The summary statistics of this dataset are very

[37]The original sample also included a total of 17,644 transactions that took place during a fortnight after the launch of the revised design on November 26, 2008, i.e. until December 10, 2008. Though all our results remain robust if we include these data, we exclude them from our analysis to account for a sufficient period for the managers to adapt to the new situation.

Figure 2.11: Price Pattern for Level-6 Keepers - Revised Design



similar to our previous sample, as shown in Table 2.15 in Appendix 2.7.2. Analogously to above, Figure 2.11 depicts the relation between total age and price for level-6 keepers from the new sample. Compared to the pattern in Figure 2.5, the formerly accentuated discontinuity between seventeen and eighteen appears to be less pronounced after the introduction of the modified design. A similar picture arises for level-7 and level-8 keepers.[38]

To analyze the impact of the reduced search costs on the birthday effect, we repeat the regression approach from above for the post-change data and test whether the predictions from Hypothesis 1 can be validated. The resulting coefficients and standard deviations are shown in Table 2.8.

Observe that all coefficients that were significant prior to the changes are also significant in the revised design. For a test whether the discontinuities are persistent, again representatively focus on the full controls-setting for level-6 keepers in column II. Note that the coefficients for the age-group dummies are of considerably smaller magnitude than in the original sample, while the intercept is of similar size. Upon entering age-group eighteen, a player now loses a total of $\beta_{age18} = -76,722$ in market value relative to a player aged "17 years and 0 days". Since a marginal day in age-group seventeen decreases the price on average by $\beta_{day17} = -649$, the aggregated day-by-day value loss during this year amounts to $-72,688$, and thus explains 94.7% of the total reduction. Though the unexplained remainder

[38]See Figure 2.14 in Appendix 2.7.2.

Table 2.8: Determinants of Price - Revised Design (OLS)

	Keeper = 6		Keeper = 7		Keeper = 8	
	I	II	III	IV	V	VI
age18	-74,979.374***	-76,722.152***	-247,552.356***	-253,063.479***	-344,138.917***	-373,922.005***
	(1,967.193)	(1,891.953)	(4,919.512)	(4,580.694)	(21,135.743)	(20,840.090)
age19	-95,377.525***	-99,023.446***	-292,993.359***	-299,006.821***	-360,970.922***	-405,873.747***
	(1,878.451)	(1,822.472)	(4,665.350)	(4,315.521)	(20,730.110)	(20,263.676)
age20	-91,044.546***	-95,773.387***	-281,258.992***	-286,345.114***	-305,034.576***	-345,697.233***
	(1,913.692)	(1,859.194)	(4,654.779)	(4,317.355)	(20,615.772)	(20,362.150)
days17	-619.094***	-648.835***	-2,149.751***	-2,230.041***	-3,059.448***	-3,146.720***
	(25.501)	(24.364)	(63.980)	(59.669)	(242.495)	(239.688)
days18	-32.728***	-51.571***	-148.830***	-150.428***	-345.155***	-223.779*
	(10.808)	(10.081)	(35.694)	(32.978)	(119.026)	(118.265)
days19	-4.102	-12.197*	-13.270	0.351	-316.720***	-51.186
	(6.808)	(6.540)	(24.389)	(22.355)	(107.828)	(99.160)
days20	5.301	1.360	-7.664	-26.162	-132.374	30.672
	(8.833)	(7.986)	(23.729)	(21.219)	(110.512)	(97.709)
tsi	19.451***	26.708***	42.352***	64.126***	55.661***	71.473***
	(0.519)	(1.474)	(1.245)	(2.358)	(2.566)	(4.189)
Intercept	77,854.031***	55,915.107***	271,900.444	147,524.418***	569,769.264***	332,985.430***
	(1,933.198)	(4,856.758)	(5,858.930)***	(10,393.075)	(23,164.350)	(31,986.616)
skills	no	yes	no	yes	no	yes
character	no	yes	no	yes	no	yes
daytime	no	yes	no	yes	no	yes
weekday	no	yes	no	yes	no	yes
R^2	0.53	0.58	0.63	0.68	0.34	0.46
N	14,873	14,873	11,942	11,942	3,480	3,480
F	910.06	307.87	1050.61	353.10	152.49	66.17

Notes: Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

of $-4,034$ still identifies a highly significant discontinuity at the eighteenth birthday (p-value: 0.0062), its magnitude has substantially decreased from 20.6% of the total decline before the change to 5.3% in the revised design. Even more intriguing, for level-7 and level-8 keepers (columns IV and VI) the prediction $\beta_{age18} = 112 \cdot \beta_{day17}$ cannot be rejected at p-values of 0.4195 and 0.1076, respectively. More precisely, the formerly strong discontinuities have vanished for these *keeping*-levels. These findings clearly indicate that the changes in the game design and the implied reduction in search costs effectively reduce the birthday effect at the eighteenth birthday considerably.

A qualitatively similar result arises where a player turns twenty.[39] However, we find no significant reduction in the discontinuity at the nineteenth birthday. If a player turns nineteen, he still experiences a strong and significant drop in value. Though surprising, this finding is in line with the newly introduced search restriction regarding the age-groups. Intuitively, if a manager enters the transfer market and conducts a combined search for

[39] In fact, in some specifications the direction of the birthday effect is even reversed if a player enters age-group twenty, i.e. we find a small increase of market value. While this does not conflict with our argument, this is surprising and most likely linked to the newly introduced age-group restriction in the search filter.

seventeen and eighteen year old players, it is likely that his next search inquiry does not again include age-group eighteen, but rather age-groups nineteen and twenty. Thus, the connection between the age-groups eighteen and nineteen is potentially not established as strongly as that between seventeen and eighteen, which would explain the finding that the discontinuity at the nineteenth birthday persists.

To further validate the impact of the reduction in search costs, we estimate a fully interacted model for both the pre- and post-design change data, where we indicate the post-change observations by the dummy variable *post* taking value 1 and 0 otherwise.[40] With this approach, we are able to test directly whether the estimates for the two samples are significantly different from each other. In the following, coefficients for the standard variables reflect the effect for the initial sample, while those for variables headed by a Δ-sign capture the relative difference for the post-design change sample. Hypothesis 2 states the predicted signs of these differences, given that it is indeed search costs that drive the birthday effect.

Hypothesis 2 *If the discontinuities are the outcome of a costly search procedure, an exogenous reduction in the search costs will lead to*

(i) a lower total decline in market value across two age-groups, i.e.

$$\Delta\beta_{age18} > 0,\ \Delta\beta_{age19} > 0,\ \textit{and}\ \Delta\beta_{age20} > 0, \qquad \textit{and}$$

(ii) an equal or increasing impact of a marginal day within an age-group, i.e.

$$\Delta\beta_{day17} \leq 0,\ \Delta\beta_{day18} \leq 0,\ \Delta\beta_{day19} \leq 0,\ \textit{and}\ \Delta\beta_{day20} \leq 0.$$

Table 2.9 depicts the results from the pooled regression. First, observe that all Δ-coefficients for the age-group dummies have a positive sign and are highly significant with exception of β_{age19} for level-6 keepers. Evidentially, and consistent with our prediction, the total decline in market value has decreased after the design change.

Regarding the second prediction from Hypothesis 2, the evidence is mixed. Almost all of the Δ-coefficients for the variables *days17-days20* turn out to be insignificant, indicating that the impact of a marginal day has not changed. Moreover, the fact that most have a negative sign implies a trend towards an increased impact of the precise age. However, for level-7

[40] See Appendix 2.7.1 for a detailed description of this approach.

Table 2.9: Comparison - Difference in Effects Post Design Change

	Keeper = 6	Keeper = 7	Keeper = 8
age18	-88,096.64***	-325,926.07***	-455,444.61***
	(2,574.46)	(8,149.90)	(30,737.96)
Δ age18_after	11,374.49***	72,862.59***	81,522.60**
	(3,195.17)	(9,349.82)	(37,154.24)
age19	-100,614.80***	-368,241.79***	-524,267.98***
	(2,512.78)	(7,788.61)	(29,974.36)
Δ age19_after	1,591.35	69,234.97***	118,394.23***
	(3,104.37)	(8,905.05)	(36,198.25)
age20	-102,967.59***	-383,271.15***	-560,486.34***
	(2,537.94)	(7,831.03)	(29,917.91)
Δ age20_after	7,194.20**	96,926.03***	214,789.11***
	(3,146.35)	(8,943.06)	(36,206.98)
days17	-624.86***	-2,614.91***	-3,128.98***
	(33.10)	(101.82)	(361.33)
Δ days17_after	-23.98	384.87***	-17.74
	(41.10)	(118.03)	(433.80)
days18	-58.32***	-255.88***	-215.21
	(13.08)	(42.56)	(157.82)
Δ days18_after	6.75	105.45*	-8.57
	(16.52)	(53.85)	(197.33)
days19	4.40	-22.16	12.92
	(10.01)	(29.60)	(119.93)
Δ days19_after	-16.60	22.51	-64.11
	(11.96)	(37.10)	(155.71)
days20	4.70	-8.29	227.80*
	(9.68)	(25.60)	(118.48)
Δ days20_after	-3.34	-17.87	-197.13
	(12.55)	(33.26)	(153.67)
tsi	26.00***	74.92***	81.54***
	(1.61)	(3.09)	(5.10)
Δ tsi_after	0.71	-10.80***	-10.07
	(2.18)	(3.89)	(6.60)
Intercept	56,549.64***	308,490.39***	554,600.70***
	(5,545.37)	(13,725.89)	(44,433.14)
Δ Intercept_after	-8,890.82	-188,817.70***	-276,287.70***
	(8,273.94)	(19,630.85)	(59,723.19)
skills	yes	yes	yes
character	yes	yes	yes
daytime	yes	yes	yes
weekday	yes	yes	yes
R^2	0.60	0.71	0.54
N	23,968	18,618	5,219
F	281.95	310.84	69.68

Notes: The coefficients before the change (only age-group in preview) are represented by the standard variables. The Δ-coefficients capture the relative difference in the impact of a variable after the change in the game design (full age in preview). The coefficients differ across the two groups, if the latter are significant. Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

keepers we find that the impact of *days17* and *days18* is significantly smaller after the design change. Yet, our findings clearly indicate that the birthday effect is mitigated though not fully explained by the reduction in search costs.

Result 3 *Search costs affect the intensity of the birthday effect but cannot solely explain its existence.*

2.5.2 Other Explanations - Heuristic Decision Making

In general, numerous other factors can affect how individuals go about making decisions. Among others, often some "rule of thumb" or heuristic is employed to simplify the procedure. Likewise, sometimes the way the different options are "framed" may lead individuals to act differently than they might otherwise. In this specific setting, among the set of attributes, age is the only one that explicitly consists of two dimensions, *years* and *days*. Thus, the managers may be lead astray by using some kind of representativeness heuristic, or base-rate fallacy, by taking the age in years as an overly informative indicator for the precise age, while neglecting the actual distribution of the variable *days*. Once they enter the profile page of a player, however, they are automatically presented with the exact age and should update the information received earlier accordingly.

Another possibility is that the managers apply some form of a sequential choice heuristic or rational shortlist method (see e.g. Manzini and Mariotti, 2006 and Zwick et al., 2003). Within such a strategy, "inferior" alternatives are sequentially eliminated in successive steps by an application of a set of criteria, until only a small set of alternatives is left from which the decision maker chooses. Translated to the situation on *HT*'s transfer market, consider a manager who is looking for a keeper-trainee. In a first step, he chooses a distinct level for the *keeping*-skill and eliminates all players that do not fulfill this criterion. Then he decides for, say, the age-group the player should belong to, and accordingly reduces the set of alternatives by elimination of all players displaying a different value for *years*, and so forth. Given this decision making procedure, it is a natural consequence that the managers compare players to the respective average within the same age-group, thereby neglecting the connection to the players from the subsequent age levels.

2.6 Discussion and Conclusion

We employ a hedonic regression approach to examine empirically to what extent managers playing the online game Hattrick incorporate available information when bidding in auctions for virtual players reflecting complex goods of multiple attribute dimensions. Using detailed field data from the game's internal transfer market, we find strong evidence that individuals inefficiently utilize substantial parts of valuable information, even if it is readily provided. Though the precise age of a player is clearly stated, the pattern of winning bids exhibits distinct discontinuities where the players grow one year older. As a consequence, the managers systematically overpay for players close to their birthday. This finding proves highly robust across all estimation approaches and if we control for experience. By exploiting a change in the game design, we analyze whether this "birthday effect" can be explained by classic search costs. We find evidence that a reduction in search costs mitigates the intensity of the documented frictions, but cannot solely explain them. We conclude that the managers disproportionately cling to the figure displayed in the variable *years*, while in turn under-weighing the information embodied in the *days* of age. Important information is not efficiently incorporated within their bidding considerations.

One implication that arises from this evidence is linked to the field of shrouded attributes and obfuscation. The basic idea in this literature is that parts of crucial information are strategically clouded or held back from consumers to hamper their individual evaluation of the product.[41] In contrast, our findings suggest that even if the full attribute vector of such a good is openly accessible, customers may simply not incorporate essential parts of it in their pricing decisions. Yet, Ellison & Ellison (2005) argue that obfuscation can take forms "*as simple as making product descriptions complicated [...] so that consumers have to examine the attributes and prices of a large number of products to know what is being offered*" (cf. p.3). The large number of similar players offered simultaneously on the transfer market and the numerous details available for each player may itself act as an "implicit obfuscation device", triggering heuristic decision making on behalf of the bidders. Even though, the neglect of profoundly important information and the magnitude of the resulting discontinuities we observe in our data remain stunning. Intuitively, excess availability of information may in fact hinder efficient information aggregation as much as does the lack of information. Our findings thus suggest that a consumer-friendly market designer or regulator should ensure that emphasis is placed on the most relevant information rather than to enforce mere information disclosure.

[41] Gabaix & Laibson (2006) show that it can be optimal for firms to cloud prices for add-on products given that at least some customers act myopic. Ellison (2005) provides evidence that primary products (e.g. printers) are often under-priced and heavily advertised to attract customers, while profits are actually generated through overpriced add-on goods (e.g. cartridges). Hotz & Xiao (2007) show that sellers can have an disincentive to disclose all relevant information on multi-attribute goods if facing heterogeneous buyers.

If even negligibly small costs such as a few "clicks" in the internet browser can suffice to affect individual bidding behavior considerably, our findings also bear important implications for the optimal design of internet auction markets. Due to a rapidly growing number of real estates, cars, and other complex goods being sold in internet auctions, people are increasingly confronted with the task to evaluate such goods by accounting for their constituent characteristics. In light of the large stakes these consumer decisions involve, obviously we would expect much more careful considerations given to the evaluation procedure. Yet, we cannot preclude that buyers of such goods are subject to a similar behavior as the managers in *HT*. While in financial markets arbitrageurs would eventually correct for systematic mispricing, in auction markets it is exactly those people who make the biggest mistakes that determine the final price. Hence, to avoid prices being systematically above fundamentals and inefficient allocations, from a social planner's perspective careful considerations should be given to the design of the market platform and the way in which relevant information on a complex good is presented to prospective buyers.

Several extensions to this research suggest themselves. When documenting systematic biases like the birthday effect, a natural next step is to ask whether these are strategically exploited by rational subjects in the market. Assuming the sellers take the demand side behavior as given, there are at least two effects we would expect to observe if the biased bidding behavior of the buyers was exploited: First, there should be a spike in the number of sale offers for players that are close to turn one year older. Second, there should be a sharp drop in askprices close *before* a player has his birthday. Intuitively, the higher the askprice, the higher is the risk that the player remains unsold. To avoid the value loss that realizes when a player turns one year older, the selling manager should rationally set a lower minimum bid, thereby increasing the probability of a successful sale. To analyze the supply side of the transfer market, in Englmaier and Schmöller (2009a), Chapter 3 of this thesis, we exploit a different sample from *HT*, where we have access to details also on failed auctions and the respective reserve prices that were set. Among others, we find a clustering of sale offers before the switch-points in these data and the median askprice is indeed substantially smaller in close proximity to the birthdays. In general, the reserve price pattern is qualitatively shaped remarkably similar to the sales price pattern.

Even though we have strong reasons to believe that the managers have serious incentives when engaging on the game's transfer market, an obvious caveat with this data is the fact that all transactions are carried out in terms of virtual currency. Therefore, it would be interesting to analyze whether similar biases as the birthday effect can be verified within real markets involving real monetary stakes. For instance, cars closely resemble the virtual players in our data in several respects. They are traded in large numbers on specialized online market platforms an their value can be decomposed into their constituent characteristics. Most importantly, and similar to the virtual players, the age of a car is commonly displayed through two dimensions, the year and month of construction or initial registration. Though buying a car constitutes a major purchase for most households and thus should be subject to profound pricing considerations, similar to the managers in *HT* people may systematically underrate the information on the precise age, i.e. the month of first registration in this case. In that case, we would expect to observe congeneric discontinuities in the price pattern for used cars where the year of first registration changes. In Englmaier and Schmöller (2009c), which constitutes Chapter 4 of this dissertation, we therefore examine a large sample of used car offers from a leading online marketplace to test for the external validity of the birthday effect. Though these data do not originate from an auction environment and naturally contain considerably more noise, the basic situation is comparable to that of the managers on *HT*'s transfer market: People are presented with a lot of details on each item and they have to estimate their valuation. We are able to document statistically significant discontinuities of considerable magnitude in the pattern of stated prices where the year of first registration changes. This indicates that an evaluation bias like the birthday effect can have real economic consequences.

Games like *HT* where people strategically interact not only attract increasing numbers of users, but also provide sufficient control to serve as novel platform for economic and psychological research. We are convinced that the vast amount of data generated by thousands of highly motivated online gamers provides a fruitful source for valuable insights and can contribute to the analysis of human decision making.[42]

[42] To our knowledge there are two other studies using data from *HT*. Ajalin et al. (2004) address the issue of betting on virtual gambles, and Trautmann and Traxler (2009) analyze whether reserve prices act as a psychological reference point for the final price.

2.7 Appendix

2.7.1 Comparing Estimates Across Samples

In Section 2.5 we have analyzed two situations, where we compare the estimates obtained from two separate samples. For both the broadband proxy approach and the before-after comparison with respect to the design change, we applied a combined regression approach to test for relative differences in the coefficients obtained for two different groups, which is presented in the following.

Consider a pooled sample containing the data for two different groups $i \in [1, 2]$. Denote the dependent variable by y and the regressors by x_1 and x_2. To test for differences between the coefficients across the two groups, we estimate the following model:

$$y = \alpha_1 + \beta_1 \cdot x_1 + \gamma_1 \cdot x_2 + \tilde{\alpha_1} \cdot g_2 + \tilde{\beta_1} \cdot x_1 \cdot g_2 + \tilde{\gamma_1} \cdot x_2 \cdot g_2 + u,$$

where the dummy $g_2 = 1$ for an entry from group 2 and $g_2 = 0$ otherwise. By this definition, the estimation model for a member of group 1 ($g_2 = 0$) is given by

$$y = \alpha_1 + \beta_1 \cdot x_1 + \gamma_1 \cdot x_2 + u.$$

Respectively, for a member of group 2 ($g_2 = 1$) we get

$$y = (\alpha_1 + \tilde{\alpha_1}) + (\beta_1 + \tilde{\beta_1}) \cdot x_1 + (\gamma_1 + \tilde{\gamma_1}) \cdot x_2 + u.$$

Thus, the above model is equivalent to estimating the separate models

$$\begin{aligned} y &= \alpha_1 + \beta_1 \cdot x_1 + \gamma_1 \cdot x_2 + u \\ y &= \alpha_2 + \beta_2 \cdot x_1 + \gamma_2 \cdot x_2 + u \end{aligned}$$

for groups 1 and 2, respectively, where

$$\begin{aligned} \alpha_2 &\equiv \alpha_1 + \tilde{\alpha_1}, \\ \beta_2 &\equiv \beta_1 + \tilde{\beta_1}, \\ \gamma_2 &\equiv \gamma_1 + \tilde{\gamma_1}. \end{aligned}$$

To use the notation from above, define $\Delta\alpha \equiv \alpha_1 \cdot g_2$, $\Delta x_1 \equiv x_1 \cdot g_2$, and $\Delta x_2 \equiv x_2 \cdot g_2$. The coefficient for, say, Δx_1 is thus given by $\tilde{\beta_1} = (\beta_2 - \beta_1)$, i.e. the difference between the estimates of the impact of x_1 on y across the two groups. Thus, we can use a standard Wald-test to verify whether $\tilde{\beta_1}$ is significantly different from zero. Suppose that was the case, then the impact of x_1 on y significantly differs across the two groups. In particular, suppose that $\beta_1 > 0$ and $\tilde{\beta_1} < 0$. This immediately implies that $\beta_1 > \beta_2$, i.e. the impact of x_1 is significantly lower for group 2 than for group 1. To sum up, this way we are able to formally test for variations in the impact of individual explanatory variables across two different groups in the sample population.

2.7.2 Additional Tables and Figures

Table 2.10: Denomination of Skill Levels

Skill Label	Integer Score	Skill Label	Integer Score
non-existent	0	brilliant	11
disastrous	1	magnificent	12
wretched	2	world class	13
poor	3	supernatural	14
weak	4	titanic	15
inadequate	5	extra-terrestrial	16
passable	6	mythical	17
solid	7	magical	18
excellent	8	utopian	19
formidable	9	divine	20
outstanding	10		

Table 2.11: Correlations

	price	years	days	totalage	keeper	tsi
price	1.00					
years	-0.31	1.00				
days	-0.02	-0.03	1.00			
totalage	-0.30	0.96	0.23	1.00		
keeper	0.81	-0.02	0.03	-0.01	1.00	
tsi	0.79	-0.04	0.02	-0.03	0.87	1.00

Table 2.12: Determinants of Price - Outlier Robust Regressions

	Keeper = 6		Keeper = 7		Keeper = 8	
	I	II	III	IV	V	VI
age18	-72,309.93***	-72,381.80***	-279,930.14***	-282,889.71***	-407,490.65***	-423,325.36***
	(1,367.20)	(1,253.00)	(3,997.42)	(3,791.25)	(19,124.15)	(18,277.88)
age19	-83,149.83***	-84,445.47***	-322,001.26***	-324,930.55***	-472,199.34***	-492,512.83***
	(1,246.24)	(1,149.38)	(3,630.25)	(3,464.38)	(18,485.81)	(17,694.24)
age20	-84,301.64***	-86,675.94***	-337,007.38***	-339,369.15***	-516,575.15***	-528,124.89***
	(1,276.17)	(1,194.12)	(3,612.00)	(3,485.17)	(19,113.05)	(18,265.72)
days17	-459.66***	-475.63***	-2,051.63***	-2,119.59***	-2,922.17***	-2,840.45***
	(15.17)	(13.95)	(42.61)	(41.18)	(212.52)	(210.03)
days18	-38.64***	-55.78***	-235.38***	-255.17***	-428.58***	-246.14*
	(14.48)	(13.32)	(44.13)	(41.82)	(139.27)	(142.406)
days19	8.72	5.05	-69.18*	-24.40	-249.29*	0.52
	(12.29)	(11.40)	(36.87)	(34.93)	(131.19)	(132.747)
days20	-10.98	4.23	-29.28	-10.41	166.11	223.24
	(12.79)	(11.72)	(35.93)	(33.90)	(154.79)	(150.79)
tsi	19.93***	22.27***	39.11***	63.84***	57.78***	81.60***
	(0.50)	(1.23)	(1.16)	(2.02)	(2.78)	(4.34)
Intercept	79,326.41***	47,632.88***	345,900.60***	282,793.02***	707,873.92***	533,250.45***
	(1,436.90)	(4,071.88)	(5,024.90)	(9,935.37)	(21,862.41)	(34,979.61)
skills	no	yes	no	yes	no	yes
character	no	yes	no	yes	no	yes
daytime	no	yes	no	yes	no	yes
weekday	no	yes	no	yes	no	yes
R^2	0.58	0.64	0.76	0.79	0.61	0.66
N	9,095	9,095	6,676	6,676	1,739	1,739
F	1,591.18	542.05	2,694.68	850.73	339.68	111.48

Notes: RREG controls for influential outliers by computing point-specific weights for the contribution of each observation to the final regression. Standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

Table 2.13: Determinants of Log-Price - Outlier Robust Regressions

	Keeper = 6		Keeper = 7		Keeper = 8	
	I	II	III	IV	V	VI
age18	-1.0392***	-1.0668***	-0.9519***	-0.9621***	-0.5631***	-0.5657***
	(0.0291)	(0.0265)	(0.0194)	(0.0184)	(0.0321)	(0.0309)
age19	-1.3026***	-1.3561***	-1.1930***	-1.2074***	-0.6768***	-0.6946***
	(0.0266)	(0.0242)	(0.0176)	(0.0168)	(0.0310)	(0.0300)
age20	-1.3470***	-1.4288***	-1.2868***	-1.3013***	-0.7614***	-0.7581***
	(0.0272)	(0.0252)	(0.0175)	(0.0169)	(0.0321)	(0.0309)
days17	-0.0054***	-0.0060***	-0.0060***	-0.0062***	-0.0038***	-0.0035***
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0004)	(0.0004)
days18	-0.0008***	-0.0013***	-0.0013***	-0.0014***	-0.0007***	-0.0005**
	(0.0003)	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
days19	0.0003	0.0001	-0.0004**	-0.0002	-0.0005**	-0.0001
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
days20	-0.0002	0.0000	-0.0002	-0.0001	0.0003	0.0003
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
tsi	0.0004***	0.0005***	0.0002***	0.0003***	0.0001***	0.0001***
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.000)
Intercept	10.8465***	10.1063***	12.4344***	12.1021***	13.3347***	13.0276***
	(0.0306)	(0.0860)	(0.0244)	(0.0483)	(0.0367)	(0.0592)
skills	no	yes	no	yes	no	yes
character	no	yes	no	yes	no	yes
daytime	no	yes	no	yes	no	yes
weekday	no	yes	no	yes	no	yes
R^2	0.49	0.58	0.71	0.74	0.58	0.63
N	9,095	9,095	6,676	6,676	1,739	1,739
F	1,095.69	418.37	1,997.47	640.98	303.89	98.65

Notes: RREG controls for influential outliers by computing point-specific weights for the contribution of each observation to the final regression. Standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

Table 2.14: Comparison High vs. Low Search Costs (Broadband-Proxy)

	Keeper = 6	Keeper = 7	Keeper = 8
age18	-85,473.20***	-319,050.98***	-453,782.55***
	(4,770.67)	(14,635.30)	(48,305.47)
Δ age18_highcost	-3,569.14	-26,256.48	-70,984.12
	(6,422.98)	(20,315.50)	(77,840.79)
age19	-100,015.27***	-365,225.79***	-528,140.34***
	(4,558.15)	(13,744.82)	(46,017.13)
Δ age19_highcost	-1,734.76	-11,342.45	-79,247.54
	(6,214.69)	(19,352.37)	(75,938.59)
age20	-102,519.90***	-380,350.72***	-556,667.51***
	(4,574.76)	(13,837.10)	(45,857.79)
Δ age20_highcost	-451.10	-11,703.69	-88,817.39
	(6,280.25)	(19,460.05)	(75,589.80)
days17	-615.43***	-2,690.00***	-3,102.64***
	(60.09)	(177.97)	(535.54)
Δ days17_highcost	10.67	64.87	-685.09
	(83.22)	(252.69)	(933.67)
days18	-104.19***	-388.76***	-473.11*
	(27.14)	(81.59)	(286.63)
Δ days18_highcost	41.38	268.06**	344.93
	(34.98)	(105.47)	(386.38)
days19	-2.98	-50.71	11.40
	(20.25)	(53.31)	(195.15)
Δ days19_highcost	29.24	17.45	286.62
	(27.66)	(74.97)	(320.33)
days20	4.77	-28.41	174.89
	(18.61)	(45.61)	(191.40)
Δ days20_highcost	0.83	16.40	90.78
	(25.99)	(66.42)	(297.77)
tsi	25.21***	82.92***	84.99***
	(3.54)	(5.68)	(7.92)
Δ tsi_highcost	-1.23	-6.85	-15.01
	(4.48)	(8.17)	(13.37)
Intercept	53,367.05***	305,848.47***	563,108.60***
	(12,010.04)	(23,358.77)	(67,588.23)
Δ Intercept_highcost	-5,013.68	13,485.71	163,019.58
	(15,239.19)	(35,034.76)	(111,455.35)
skills	yes	yes	yes
character	yes	yes	yes
daytime	yes	yes	yes
weekday	yes	yes	yes
R^2	0.61	0.75	0.66
N	5,567	4,387	1,205
F	59.34	84.38	25.90

Notes: Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

Explanation: The proxy $prx\_sc \in [0,1]$ ranks each country relative to the high-speed internet subscriber rate of Denmark, which has the highest number of broadband users (36 percent) among the countries represented in HATTRICK. Values close to 1 indicate countries with low search costs. The regression compares the coefficients estimated over the group with high search costs ($prx\_sc < 0.40$) to the coefficients estimated over the group with low search costs ($prx\_sc > 0.61$), both accounting for roughly one third of the total sample. The result for low search costs are represented by the standard variables, while the Δ-coefficients capture the relative difference in the estimates for the high search cost group. Note that all Δ-coefficients for the age-group dummies are negative, while all coefficients for Δ*days17*-Δ*days20* except one have a positive sign. If anything, this trend point towards a more pronounced birthday effect, since the managers with high search costs seem to give even less weight to the exact age.

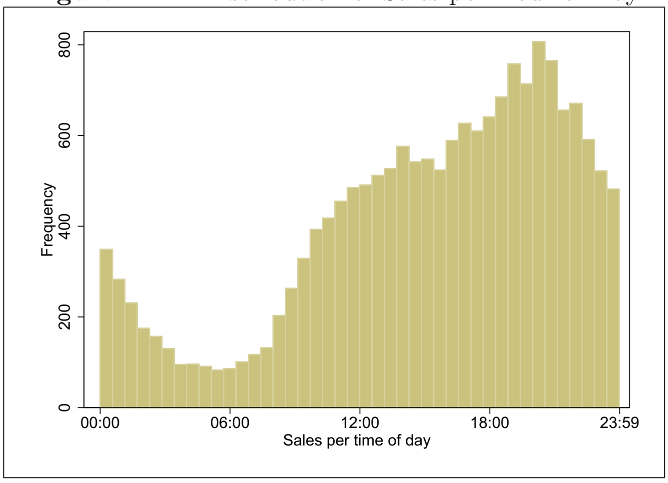
Table 2.15: Summary Statistics - Revised Design

Panel A. Overview					Panel B. Price for Age-Skill-Combinations					
Variable	Obs.	Mean	Min.	Max.	Skill	Age	Obs.	Mean	Min.	Max.
price	30,295	149,843	1,000	1,500,000	6	17	2,934	85,358	4,000	306,000
years	30,295	19	17	20		18	3,376	40,226	1,000	140,000
days	30,295	55	0	111		19	4,225	20,999	1,000	94,000
totalage	30,295	238	3	447		20	4,338	26,365	1,000	120,000
total skill index	30,295	2,900	450	8,530						
keeping	30,295	7	6	8	7	17	2,871	314,378	79,000	1,000,000
playmaking	30,295	1	1	4		18	2,493	158,222	32,000	450,000
scoring	30,295	1	1	4		19	3,089	119,046	24,000	349,000
passing	30,295	1	1	6		20	3,489	131,277	36,000	375,000
winger	30,295	1	1	5						
defending	30,295	1	1	6	8	17	702	637,780	298,000	1,500,000
setpieces	30,295	2	1	20		18	926	489,328	230,000	1,016,000
stamina	30,295	5	1	9		19	970	470,337	207,000	950,000
leadership	30,295	4	1	7		20	882	533,230	206,000	1,020,000
wage	30,295	2,159	790	4,210						
form	30,295	6	1	8						
player experience	30,295	1	1	5						

Panel C. Prices per Agegroup and Skilllevel

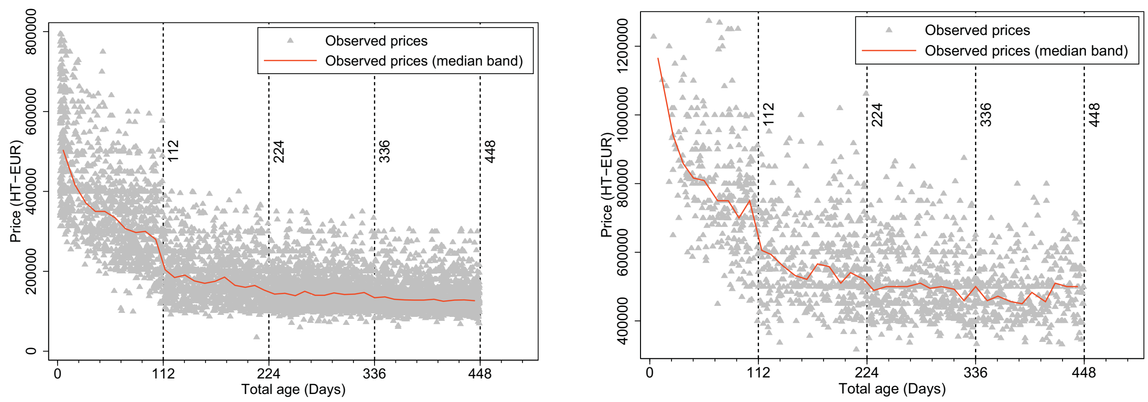
	Variable	Value	Obs.	Percent	Mean	Min.	Max.
prices by age	years	17	6,507	21.48	246,003	4,000	1,500,000
		18	6,795	22.43	144,719	1,000	1,016,000
		19	8,284	27.34	110,174	1,000	950,000
		20	8,709	28.75	119,728	1,000	1,020,000
prices by skill-level	keeping	6	14,873	49.09	39,625	1,000	306,000
		7	11,942	39.42	177,758	24,000	1,000,000
		8	3,480	11.49	525,108	206,000	1,500,000

Figure 2.12: Distribution of Sales per Hour of Day



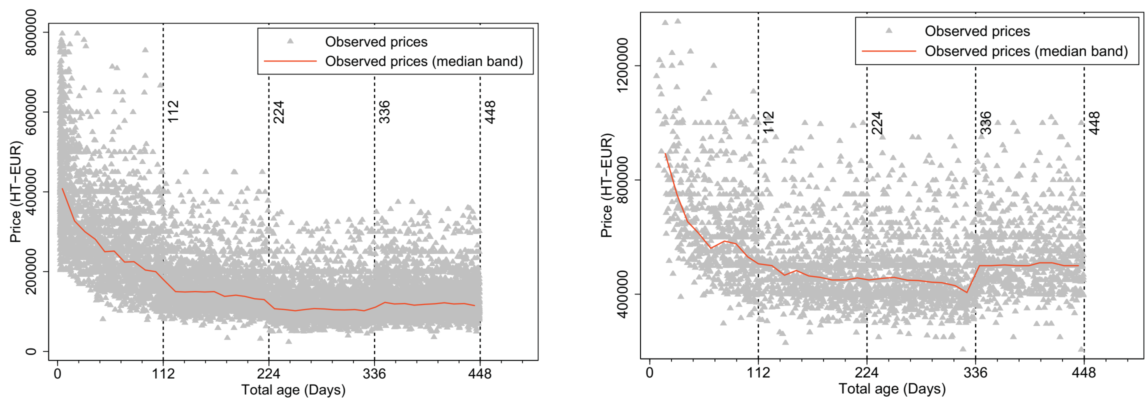
Notes: During a typical day, the time of the auction deadlines is approximately similarly distributed as the number of simultaneously logged-on users. During the early morning hours CET we observe the lowest traffic with roughly 7,000 online users at its minimum, while during the peak-periods between 5:30 p.m. and 10 p.m. levels up to 75,000 simultaneous online users are reached.

Figure 2.13: Price Pattern for Level-7 and Level-8 Keepers (Original Sample)



(a) Price Pattern for Level-7 Keepers (b) Price Pattern for Level-8 Keepers

Figure 2.14: Price Pattern for Level-7 and Level-8 Keepers (Post-Change Sample)



(a) Price Pattern for Level-7 Keepers (b) Price Pattern for Level-8 Keepers

Chapter 3

Determinants and Effects of Reserve Prices in Hattrick Auctions

3.1 Introduction

In a situation where buyers' willingness to pay is private information or the identity of the highest value buyer is unknown, an auction, instead of posting or negotiating a price, can be an efficient allocation mechanism. Hence, auctions have been used in a wide array of fields like art sale, real estate, or the allocation of spectrum rights. With the ascent of the internet, auctions have exceedingly gained popularity on platforms such as *eBay.com*, *amazon.com*, or *eBid.com*. Alone *eBay.com* is present in 39 markets and in 2007 approximately 84 million active users worldwide sold items on eBay trading platforms for nearly \$60 billion, i.e. *eBay.com* users worldwide trade more than \$1,900 worth of goods on the site every second.[1] The rise of internet auctions led to an enhanced availability of large data sets. Based on these, numerous empirical studies have addressed positive questions ("Is observed bidding consistent with Bayesian Nash Equilibrium (BNE)?", "Is there evidence of buyer risk aversion?") and normative issues (recovering the value distribution, identifying the optimal auction, then simulating the effects of design changes).

From various empirical, e.g. Lucking-Reiley (2000), or theoretical, e.g. Myerson (1981), Riley and Samuleson (1981), or Bulow and Roberts (1989), studies it is obvious that reserve prices (public minimum bids) are an important strategic design element in most auction environments.[2] Using a hand-collected data set of 6,258 auctions of virtual football players

[1] Source: http://news.ebay.com/about.cfm

[2] Rosenkranz and Schmitz (2007) extend the analysis to non-standard reference point dependent preferences and show whether and how reserve prices perceived as reference points affect bidding behavior.

traded in English auctions on *hattrick.org* we are able to address both positive as well as normative issues. We analyze how reserve prices are set and where they deviate from theoretically predicted patterns but we also perform the counterfactual exercise and show how much expected revenue is actually missed by setting suboptimal reserve prices.

The online game Hattrick (*HT*) is the world's largest online football manager game with almost one million participants. Every day about 40,000 virtual players are traded on the *HT* transfer market. By design, these trades take place in a highly controlled environment including a standardized duration for each auction, a fixed mode of how players on sale are presented, and no risk of default. Sellers are however free to choose a non-negative reserve price. Unlike many other online auction platforms, in *HT* there is no relation between the minimum bid and the transaction fees a seller is charged, which could bias individuals in their choice of a reserve price. Moreover, when *HT* players are on the market, all relevant information concerning their quality becomes publicly available. That is, there is no information asymmetry between buyers and sellers and hence no scope for a winner's curse[3] and reserve prices contain no quality signal. Thus, for the bidders the auction takes place in an independent private value (IPV) context where individual valuations are determined by idiosyncratic shocks to a common and publicly observable value. In the next section we describe the *HT* auction market in detail and explain how success in the game crucially depends on profitably trading virtual players.

From the classic contributions in auction theory, e.g. Myerson (1981), Riley and Samuleson (1981), we know that the optimal reserve price in an independent private value environment is a continuous function of the hazard rate of the distribution of valuations of buyers and does not depend on the number of potential bidders.[4] In the reserve price patterns in our data, we find both evidence for very sophisticated and boundedly rational behavior of sellers. We show that reserve prices are predicted, qualitatively and quantitatively, by the same observable characteristics that predict sales prices.

[3]The winner's curse refers to the fact that the bidder with the highest estimate of quality is likely to hold a too positive view of the true but unobservable quality (common value), which then forms the basis of private valuations of the product.

[4]Levin and Smith (1994) and Levin and Smith (1996) analyze alternative models with endogenous (and costly) entry decisions prior to the bidding stage. In their setting, the number of bidders and the optimal reserve price actually covary, implying that under IPV small or no reserve prices are optimal as this attracts more bidders. Conversely, they argue that a positive reserve price is useful in reducing the number of bidders in a common value auction, since the winner's curse is the worse the more bidders are participating. However, in our *HT* environment costs of entry (bid preparation, information gathering) are negligible and hence we treat the number of bidders as exogenously given.

In particular, we find that *reserve prices exhibit the "birthday effect"* that has been documented for sales prices in the previous chapter. A player's value in the game decreases, ceteris paribus, continuously with his age measured in days as it becomes harder and harder to improve his skills by training. In Chapter 2 we show a very strong drop in sales prices just on a player's birthday, indicating that buyers in *HT* give too much weight to the age of a player measured in years as opposed to his age measured in days, though the latter is also plainly visible to all buyers free of cost as can be seen in Figure 3.1. Our analysis clearly shows that the birthday effect is also present with respect to reserve prices. Further examination of the data indicates that the presence of the birthday effect is not (only) due to the fact that sellers fall prey to the same information under-usage as the buyers, but because at least a substantial fraction of sellers tries to strategically exploit this bias of the demand side. We find a clustering of sale offers just before players' birthdays, indicating that sellers rationally want to sell players before they drop in value on their birthday. Furthermore, a sharp drop of median reserve prices immediately before the birthday indicates that sellers anticipate the immanent drop in market value. Hence, they want to make sure that the player is actually sold where a (too) high reserve price might endanger this.

We run hedonic regressions with the sales price and with the reserve price as dependent variables and identical sets of explanatory variables. As stated above, all of them have qualitatively similar effects – with one notable exception: we have a good proxy whether a player on the transfer market had been acquired by the seller previously or whether he was promoted (basically for free) from the seller's own youth team. We find that, ceteris paribus, sellers set *significantly higher reserve prices* (by about 23% of the mean) for players they have bought as compared to players they promoted internally. In contrast to that, *sales prices are significantly lower* (17% of the unconditional mean price) for previously traded players unconditional on a successful sale. In Section 3.3 we discuss in detail why the negative effect we observe in sales price patterns is what we would expect from rational actors. The positive reserve price premium we find for previously traded players is more in line with the *sunk cost fallacy*, for example due to loss aversion with respect to the previous selling price, leading to an entitlement effect.

Finally, we find that *reserve prices are too clustered* as to be compatible with fully rational behavior. In particular, the reserve price pattern spikes dramatically at multiples of €50,000, and also suggests a lower scale clustering at multiples of €5,000. Moreover, a large fraction of the sellers (18%) sets a reserve price of zero. We interpret this as evidence for sellers using a round number heuristic in setting reserve prices, thereby making inefficient use of the reserve price instrument. However, we are able to do more. From observing the sales prices, in the English auction resembling the second highest bidder's valuation, we are able to estimate the underlying distributions of valuations, $F(v)$, calculate the optimal reserve prices given $F(v)$, and calculate the share of expected revenue lost at the actual levels as compared to the situation with optimal reserve prices.

Internet auction data have been widely used. Lucking-Reiley (2000) presents data from a comprehensive study of 142 different internet auction sites and describes the transaction volumes, the types of auction mechanisms used, the types of goods auctioned, and the business models employed at the various sites. Bajari and Hortacsu (2003) go deeper in their analysis and show that in their sample of coin auctions on *eBay.com*, reserve prices are set below the book value of their coins. In contrast, we find that while many reserve prices are set very low (or even at zero), a substantial fraction is set very high. As we do, Bajari and Hortacsu (2003) estimate the value distribution $F(v)$ from the observed bids and use it to evaluate the effect of alternative reserve prices. However, they do not solve for the optimal reserve price as a benchmark as they cannot rule out common value elements in their data.

The evidence on whether or not reserve prices are revenue enhancing is somewhat mixed. Ariely and Simonson (2003), Kamins et al. (2004), and Lucking-Reiley et al. (2007) show in field experiments that selling prices increase in the reserve price. Reiley (2006) documents evidence from a field experiment on ebay that reserve prices reduce the number of bidders, the probability of an actual sale, and the unconditional expected revenue. However, the expected revenue conditional on a sale in fact increases. On the contrary, Bajari and Hortacsu (2003) and Hoppe and Sadrieh (2007) find in their field study and field experiment respectively no positive effects of reserve prices on selling prices. Finally, Simonsohn et al. (2008) take a different tack. They document that eBay bidders prefer auctions with more bids, hence sellers have an incentive to set low reserve prices.

Of particular interest to us is the study by Trautmann and Traxler (2009), which also uses data from auctions of players in *HT*. Their focus is to separate two potential channels how reserve prices might affect selling prices: A reference point or anchoring effect as suggested in Rosenkranz and Schmitz (2007) and a standard rent appropriation effect that stems from reserve prices forcing the winning bidder to pay more than the second highest bidders valuation. Trautmann and Traxler (2009) find a positive effect of reserve prices but they find no evidence that any of these higher prices stem from a reference point effect but rather can be accounted for by rent appropriation.

The sunk cost fallacy, sometimes referred to as "irrational escalation of commitment", has been studied by psychologists for decades, e.g. Staw (1976), Arkes and Blumer (1985), or Bazerman (1986). In most recent discussions in economics, it has been related to prospect theory (Kahneman and Tversky, 1979), specifically to a reference point and loss aversion. In many contexts, the presence of a sunk cost fallacy implies entitlements. There is ample evidence for the entitlement effect and the resulting Willingness-to-Pay/Willingness-to-Accept gap in controlled experiments[5], though recent work by Plott and Zeiler (2005; 2007) is critical with respect to the validity of the endowment and entitlement effects outside the lab. Our study indicates a strong and persistent entitlement effect in a very competitive natural setting. All our results are robust if we control for the experience of the sellers in our sample.

Genesove and Mayer (2001) analyze seller behavior in the Boston residential real estate market using proprietary panel data. Sellers whose condominium's expected selling price falls below the original purchase price due to an aggregate market downturn tend to set asking prices well above the expected price level. They argue that this unwillingness to accept market prices for property in the down part of the market cycle could stem from loss aversion on behalf of the sellers. However, in our sample there are no losses caused by business cycle swings, since all data was collected within only a fortnight. The setting studied in Genesove and Mayer (2001) involves bargaining, where the final price can fall below the initial asking price of the seller, which in addition not necessarily reflects his reserve price. It is also reasonable to assume that the evaluation of a condominium may involve substantial search costs, whereas all relevant information on *HT*'s virtual players is readily available, highly standardized, and thus easily comparable. Hence, while we observe similar behavior of sellers, the motivations behind it may differ considerably.

[5]For references see e.g. Thaler (1980), Knetsch (1989), Hanemann (1991), Shogren et al. (1994), Casey (1995), or Carmon and Ariely (2000).

Clustering of stock prices at integers, i.e. that limit sell orders and also prices tend to be rounded to whole numbers rather than displaying fractions, has been documented by Niederhoffer (1965), Harris (1991), and Sonnemans (2006), who also provides an overview of related studies on price clustering in stock markets, discussing possible explanations. Benartzi and Thaler (2007) show that a round number heuristic seems to be important in determining savings choices, too.

The remainder of the paper is structured as follows. Section 3.2 describes the structure of the data and the relevant details of *HT*. Section 3.3 presents our empirical analysis and results. In Section 3.4 we estimate the share of expected revenue lost relative to the optimum and Section 3.5 concludes. An Appendix collects additional Tables and Figures.

3.2 Data Description

3.2.1 Institutional Background about Hattrick

Hattrick, founded in 1997, is a browser-based free online football manager game with almost one million registered users, henceforth referred to as "managers".[6] The basic concept of the game is to manage your own virtual football club, which consists of virtual players that are represented by a multi-dimensional vector of attributes. A team plays at least one weekly game in a national league system against teams coached by other managers. In *HT*, a season, or an in-game year, lasts for 112 real-time days. The outcome of matches is determined by random simulation on the basis of the chosen strategies of the opponents, skills of the virtual players and other factors that determine the probabilities to win. The tasks for a manager to lead his team to success are numerous, ranging from decisions on match tactics and line-ups, over hiring team staff like doctors and co-trainers, over "drafting" a new player from the team's youth squad and either selling, keeping, or firing him as needed, to monitoring the team's training program. Many managers complete all of these tasks almost on a daily basis.

The sportive aspect is but one of the supporting pillars of the game. Maintaining a virtual club requires the managers also to develop a solid financing scheme. The most important source of (in-game) revenue for a manager is successfully trading players on the *HT* transfer

[6] We refer to human users as "managers", while using the term "player" to address virtual football players.

market. Most managers follow a "train and trade"-strategy which first ensures the improvement of quality of their own virtual players by choosing a training scheme and then profitably selling them to other managers. Since the proceeds from player sales are the major source of income in *HT*, the transfer market provides strong incentives for the participants in this open-ended manager game.

3.2.2 Goods: The Virtual Players

A virtual player consists of about thirty dimensions. Figure 3.1 depicts the typical profile interface for a player, including the set of his attributes and the corresponding auction details. The most important attributes are the eight abilities displayed in the lower middle of the profile which we denote as his "skills".[7] While *stamina* and *set-pieces* are general skills, the remaining six - *playmaking, winger, scoring, keeping, passing,* and *defending* - determine a player's suitability to play in certain positions in the line-up. For instance, a player with his best skill being *keeping* is rationally classified as goalie. In the terminology of the game, skill levels are denoted as adjectives. To simplify the notation, we use integer values to address them, e.g. "passable" corresponds to score 6 of 20.

Figure 3.1: Virtual Player Profile

Notes: Next to the player attributes, on the lower right all information regarding the auction details are displayed. The seller of this keeper set him on the transfer market during the evening on February 01, 2008 (i.e. exactly 72 hours before the deadline displayed) at a reserve price of €25,000. (Source: www.hattrick.org)

[7] Table 2.10 in Chapter 2 shows the detailed ranking of all skill levels, which can also be found in the game's manual. To all other attributes, we will refer to as "characteristics". For a classification of the values each attribute can take, refer to Table 3.1.

From the set of player attributes, only these eight skills can be actively improved by training. However, it takes several weeks for an individual player to increase by a full skill level and only a single skill can be trained at a time. Hence, managers have to specialize in training only one specific skill, say *keeping*. As soon as such a keeper-trainee surpasses the threshold for a skill-up in this skill, the manager can profitably sell him to another manager and assign the free training slot to a new (and younger) trainee, which he can either acquire on the transfer market or promote directly from his own youth team.[8] The proceeds from the sales are in turn used to finance the club.

The value of a player crucially depends on two main factors, his current skill levels and his age. The first determine the strength a player adds to a team if he is currently lined-up in a certain position for a match, which is independent of a player's age. For the purpose of this paper, we refer to this component as his "Consumption Value" (CV). For instance, a keepers' CV is almost exclusively driven by the goalkeeping-skill. However, the second main channel of influence for the value of (young) players arises due to the fact that age is a key determinant for training effectiveness. In *HT*, the marginal skill-improvement from training declines with the age of a player. The younger a player, the more he benefits ceteris paribus from training and the faster he advances to a higher skill level, which in turn increases his CV. As a consequence, a viable training strategy necessarily requires rather young players, since they have the highest innate potential for further skill development, or "Advancement Potential Value" (APV) as we label it. Importantly, the marginal effect of training is otherwise homogeneous for all virtual players, i.e. there exists nothing like a talent-attribute capturing the potential for skill-improvements. It does to some extent depend on the ability of the club's trainer, and the training intensity chosen by the manager, where the latter two give rise to variation in private valuations.

Ceteris paribus, a player who is just a few days younger than another should not be worth much more, since the difference in their APV is minimal. This holds true irrespectively of whether one already turned a year older while the other's birthday lies just ahead. Yet, in the dataset analyzed in Chapter 2 we find that the observed sales price pattern exhibits strong discontinuities at the birthdays of the virtual players, which we refer to as the "birthday

[8] Each manager can promote one player from the youth team each week, whose attributes are determined randomly with a high probability of low skill levels. Age is also randomly assigned on the interval 17 to 19 years. All descriptions of the game are based on the set of rules and institutions that were in place during the collection period of our data set.

effect". The buyers overreact to the informational content of the age-group indicator *years*, while disregarding the finer information on a player's age attribute as conveyed through the *days*, even though the precise age of a player is explicitly stated in the form "X years and Y days" on his profile page (see Figure 3.1). Naturally, this raises the question of whether the reserve prices set by the sellers pick up the birthday effect, or whether sellers react strategically to the documented buyers' behavior. Along with the individual values they can take in the game, Table 3.1 provides an overview of the player attributes and other variables we employ for our analysis.[9]

Table 3.1: List of Variables

	Variable	Description	Range
Player attributes	years	Age in years (1 HT-year ≡ 112 real-time days)	17+
	days	Age in days (1 HT-day ≡ 1 real-time day)	{0,..,111}
	totalage	Precise age of a player in day units (normalized)	{0,..,335}
	days17-days20	Interaction term of days and age-group dummies	{0,..,111}
	form	Current form of player	{0,..,8}
	total skill index	Noisy indicator of overall quality of player	$\mathbb{N}_+$
	wage	Salary (exogenous; in virtual Euro)	$\mathbb{N}_+$
	keeper	Playing skill, position specific	{0,..,20}
	playmaking	Playing skill, position specific	{0,..,20}
	winger	Playing skill, position specific	{0,..,20}
	scoring	Playing skill, position specific	{0,..,20}
	passing	Playing skill, position specific	{0,..,20}
	defense	Playing skill, position specific	{0,..,20}
	setpieces	Playing skill for all player types	{0,..,20}
	stamina	Playing skill for all player types	{0,..,20}
	gentleness	High value if agreeable (ascending order)	{0,..,20}
	aggression	Low value if player aggressive (descending order)	{0,..,20}
	honesty	High value if honest (ascending order)	{0,..,20}
	plrexp	Experience of player	{0,..,20}
	ldrshp	Leadership qualities of player	{0,..,7}
Auction Data	askprice	Reservation price set by the seller	$\mathbb{N}_+$
	price	Auction end price paid by winning bidder	$\mathbb{N}_+$
	dtime	Time of deadline	hh.mm.ss
	dday	Day of deadline	dd.mm.yy
	sellerxp	Proxy for seller experience by relative country ranking	$[0-1]$
Dummy variables (1=yes, 0=no)	age17 - age19	Dummy for age-group	{0,1}
	peakhour	Did auction end during peak hour (5:30 p.m. - 10:00 p.m.)	{0,1}
	mon - sun	On which weekday did the auction end?	{0,1}
	acquired	Proxy for previous sale (player countryID = seller countryID)	{0,1}
	d_ask	Was the reserve price different from zero?	{0,1}
	d_sold	Was the player successfully sold?	{0,1}

Among the remaining characteristics, a player's *total skill index* (*tsi*), which represents a noisy measure for his overall abilities, is the one most likely to have a (positive) influence on market value. To see this, note that *HT* calculates the skill-levels as real numbers including hidden decimal places, the so-called "sub-skills", while the player profile only displays the adjective reflecting the current integer value for each skill. With each training a player receives, the trained skill increases by a marginal increment (which is declining in age), and

[9] In the following, we use italics to denote the variables from our sample.

so does the *tsi*. While also correlated to other attributes (e.g. *form*), the *tsi* score thus constitutes a noisy signal for the sub-skills of a player, i.e. for how close he is to reach the next higher level in one of his skills. Though a complete description of all characteristics is beyond the scope of this paper, it should be noted that for our empirical estimation we control for the full vector of attributes.

3.2.3 Transactions: The Transfer Market

With an average of about 40,000 players offered for sale each single day, the transfer market in *HT* has a remarkable trading volume.[10] The selling mechanism implemented on the transfer market is an English ascending open bid auction. To sell a player, managers can specify a non-negative reserve price and submit an irreversible sell order by clicking a button. Each auction ends exactly 72 hours from submission, but the deadline is automatically extended by 3 minutes if a bid is placed within 3 minutes to the deadline. This continues until all bidders but one retire.[11] Importantly, all relevant information concerning a players quality - that is the full attribute vector - becomes publicly available once a player is offered on the transfer market. Hence, at the time of sale there is no information asymmetry between buyers and sellers and, for that matter, the econometrician.

Since all players on sale are displayed in the same standardized way as shown in Figure 3.1, the sellers have no possibility to affect the way how an individual player is presented to potential buyers. Except for the timing of the sell order, which determines the auction deadline, this leaves a seller with a single dimension of choice: The reserve price. Thus, it is reasonable to argue that careful considerations should be given to the utilization of this remaining instrument of potential influence on the auction outcome.

The transfer market provides a search tool, which allows the managers to filter for various player attributes like age-group, current bid, and up to four playing skills at desired levels. The inquiry returns a list of offered players matching the selected filter, where an abstract of their main characteristics is displayed (see Figures 2.1 and 2.2 in Chapter 2).

[10]Source: http://www.databased.at/HT/htpe

[11]Given the reserve price is set below the second highest bidder's valuation, the transfer price will equal the second-highest bid plus one discrete increment, i.e. the format is strategically equivalent to a sealed-bid-second-price auction. For reference on the effects of the employed ending rule on bidding behavior see e.g. Roth and Ockenfels (2002) and Ariely et al. (2005).

To submit a bid for a player, a prospective buyer manager must enter his profile. Each bid must at least be equal to the reserve price or above the current highest bid, respectively. Placed bids are binding and irreversible. After the auction ends, the player is automatically transferred to the winning manager's team and the seller receives the winning bid net of some small fee.[12] If a player received no bid, the auction fails and he stays with the seller.

3.2.4 Sample Selection and Data Description

Our main interest in this paper is to study the determinants and effects of reserve prices. For this purpose, we collected all publicly available information on 6,258 virtual players offered for sale on the *HT* transfer market between November 18, 2007 and December 02, 2007.[13] The sample considers the specific subgroup of keepers aged between seventeen - the youngest age possible in the game - and nineteen years, all with an identical *keeping*-skill of score 6 out of 20, i.e. "level-6 keepers". The age criterion is motivated by the facts that (i) the APV is most important at young ages and (ii) young players are heavily traded. Regarding the focus on keepers, note that the values of field player types depend not only a combination of several skills but also on a other factors, e.g. the chosen match tactics. In effect, individual skills can receive quite different weights in the evaluations across managers, making it hard to measure the impact of a specific attribute on the price, or, as in our case, the reserve price. In contrast to that, for keepers by far the most influential skill unambiguously is the *keeping*-skill, which determines their value to the largest extent. By holding the keeping-skill constant at a score of 6, which accounts for the thickest market segment of players in this category, we effectively suppress the impact of variations in the skill-dependent CV on the observed reserve prices. We are thus able to identify influential factors for the players' APVs in the sample, which crucially depend on their age, or more precisely, their precise age including the days.[14]

Next to all relevant player and auction characteristics, we also collected information on the sellers. For a subsample of 2,411 auctions we are able to construct a proxy for seller

[12]These transaction costs are negligibly small and do not affect any of our results.

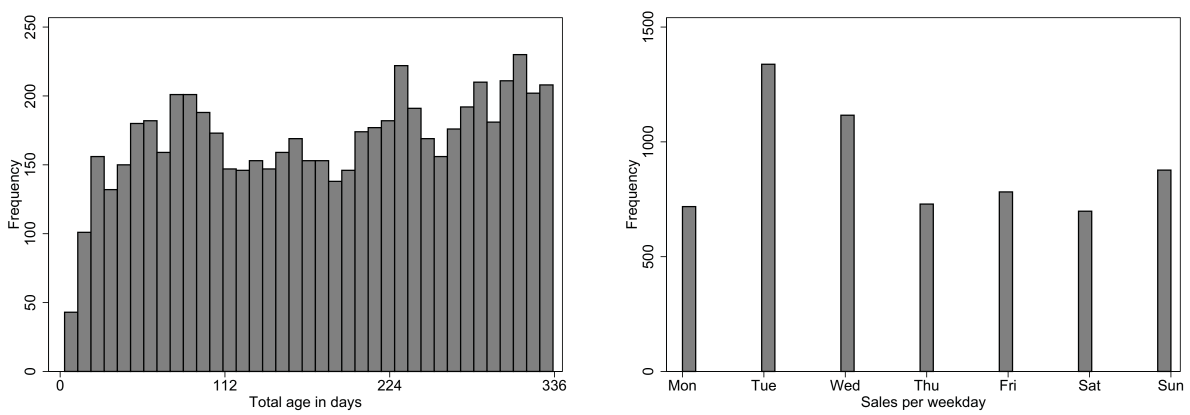
[13]From an initial sample of 6,460 players in the relevant skill- and age-group, we excluded 4 players that play for their respective home country's national team, 66 players with reserve prices identified as outliers by Grubbs' test (Grubbs, 1969) and 132 players that were injured at the time of the auction.

[14]The motivation for these selection criteria follows the lines of Chapter 2. Though many variables appear in both datasets, the present sample differs in some respects. In particular, for this study we additionally collected records of the reserve prices and more details on the individual sellers.

experience ($sellerxp \in (0, 1]$), where we use the information on how a manager ranks relative to all other managers within a given country. We argue that a higher ranking within a country is a good indicator for being more experienced as it can only be achieved by playing the game for a long period of time and/or being very successful quickly, which should to a large degree be correlated with having routine playing the game.

Furthermore, we use the information whether a player plays abroad or not, indicated by a 20% bonus on his wage,[15] as a proxy whether he had been previously traded (*acquired* = 1), or whether he is a "fresh" player from the seller's own youth team (*acquried* = 0). Fresh players never receive this 20% playing-abroad bonus on their wages, as e.g. German teams always produce fresh German players. As in roughly 85% of all trades buyer and seller are not from the same country, our potential mistake from missing trades within a country is small and this 20% bonus is a good proxy to discriminate between fresh and previously traded players.

Figure 3.2: Distributions of Age and Sales



(a) Distribution of total age (b) Distribution of sales per weekday

Figure 3.2a depicts the age distribution for the players in our sample, indicating a roughly balanced distribution for *days* within each age-group. The distribution of sales per weekday is shown in Figure 3.2b. On Tuesdays and Wednesdays we observe spikes in the number of sold players. The reason for this is that new players can be "drafted" each Saturday after an weekly update and often are immediately offered for sale, which explains the number of deadlines expiring on Tuesdays and Wednesdays being accordingly higher. In our regressions, we use dummies (*mon-sun*) to control for possible effects of the auction end

[15]Note that a player's wage is exogenously fixed by *HT* and cannot be influenced by the managers.

day and additionally also include a dummy (*peakhour*) to indicate whether a player was sold between 5:30 p.m. and 10 p.m., where the highest numbers of simultaneous online users are reached and most auctions expire.[16]

Table 3.2 shows the summary statistics of the most important variables in our sample. As shown in Panel A, the mean reserve price (*askprice*) in our sample was €77,537, but levels as high as €579,000 were reached.[17] Note that the final prices fall into a comparable range, indicating that by and large the reserve prices were not set beside the point. Since the age of a player is displayed in the form "X years and Y days" on his profile-page, the variable *years* defines his age-group and *days* discloses information on his precise age, or equivalently, the distance to his next birthday. The constructed measure *totalage* $\in [3, 335]$ displays the total age of a player in day units and thus combines the information contained in both age variables, where we normalize *totalage* $\equiv 112 \cdot (years - 17) + days$, using the fact that a year in *HT* is normalized to 112 days.[18]

Table 3.2: Summary Statistics

Panel A.				Panel B.			
Variable	Mean	Min	Max		Variable Value	Frequency	Percent
askprice	77,537	0	579,000	age distribution	(years = 17)	1,886	30.14
price[a]	81,459	0	634,000		(years = 18)	1,935	30.92
years	18	17	19		(years = 19)	2,437	38.94
days	59	0	111				
totalage[b]	181	3	335	fresh players	(acquired = 0)	4,253	67.96
total skill index	1,994	650	3,240	purchased players	(acquired = 1)	2,005	32.04
wage	1,884	770	2,676				
form	6	1	8	reserve price	(askprice > 0)	5,108	81.62
stamina	3	1	9	no reserve price	(askprice = 0)	1,150	18.38
passing	1	1	4				
playmaking	1	1	3	successful trades	(price > 0)	4,743	75.79
scoring	1	1	3	players unsold	(price = 0)	1,515	24.21
winger	1	1	4				
setpieces	2	1	7	sold at reserve price	(price=askprice)	756	15.94
defense	1	1	4	sold above reserve price	(price>askprice)	3,987	84.06

a. A Price of zero indicates a failed auction. The minimum price among all successful trades was €19,000.
b. The variable *totalage* $= 112 \cdot (years - 17) + days$ displays a players precise age in day units. The minimum value of *totalage* at 3 reflects age "17 years and 3 days" and the maximum value at 335 equals "19 years and 111 days".

Panel B provides some frequency statistics of our data. Note that all age-groups are roughly equally represented, with a slight majority of players aged nineteen. According to our wage-bonus proxy (indicated by the dummy *acquired*), in 32% of all auctions the sellers offered players they previously bought themselves on the market. 68% of the times, a player

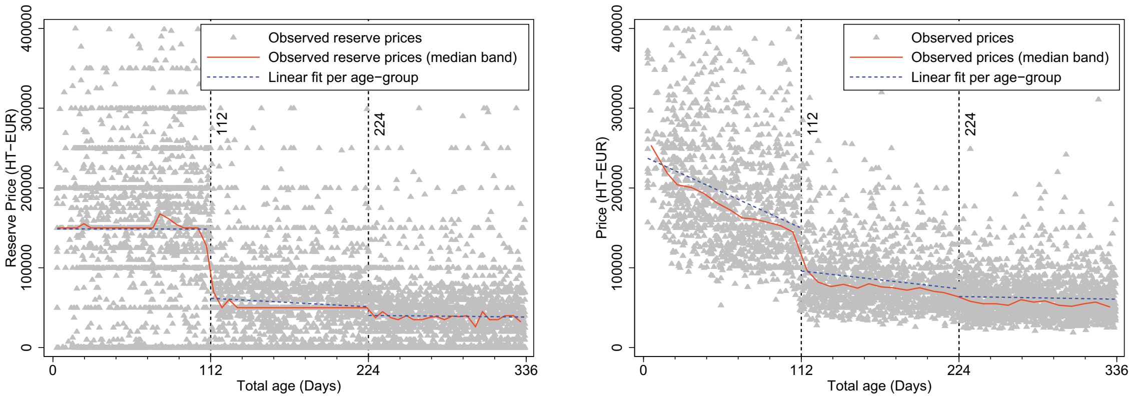
[16]See Figure 3.9 in Appendix 3.6.
[17]All monetary values are denoted in units of virtual *HT*-Euros.
[18]Note that since each auction lasts for 3 days, *totalage* has its minimum at 3, or "17 years 3 days". The maximum value of *totalage* at 335 reflects the age "19 years 111 days".

promoted from the own youth squad of the seller, i.e. a "fresh" player, was auctioned off. For 5,108 players in our sample the seller fixed a strictly positive reserve price. Surprisingly, this fact already establishes that more than 18% of the sellers did not make use of the possibility to set a reserve price.[19] Of all players, 4,743 were sold and in 756 cases the trade took place at a price equal to the minimum bid (single bidder case).

3.3 Analysis and Results

In light of the "birthday effect" on the demand side of *HT*'s transfer market, which is illustrated in Figure 3.3b analogously to Chapter 2 for the present sample, it seems a natural point to start our analysis of the driving forces in the formation of reserve prices by examining their relation to a player's age.

Figure 3.3: Relation of Reserve Prices and Sales Prices to Total Age



(a) Pattern of Reserve Prices (b) Pattern of Sales Prices

Notes: The dashed lines at 112 and 224 mark the points where the players turn eighteen and nineteen, respectively.

Intuitively, the variable *totalage* captures all available information on the players' age attribute. As we argue above, and as implied by the nature of the game's training algorithm, the value of a player should ceteris paribus decline continuously as *totalage* increases. It seems plausible to reckon that this fact should also be reflected in the reserve prices.

However, Figure 3.3a already reveals that the relation between reserve price and total age is not smooth but exhibits large discontinuities where the players enter the next higher

[19] In line with Simonsohn et al. (2008), it is possible that these managers may want to maximize entry and the number of bidders in the auction by not screening out any low-value bidders. As we will show in Section 3.4, however, from perspective of expected auction revenue they forego potential profits.

age-group, indicating that the birthday effect also persists for reserve prices. Since most managers in *HT* alternate between both roles, it does not seem too surprising to find similar behavioral patterns for buyers and sellers. At a first glance, it stands to reason that also the sellers inefficiently utilize the finer information on the age attribute as conveyed through the *days* of age. Another possibility is that the bias on the demand side actually triggers the observed choice of reserve prices, and what we observe is the result of at least some sellers following strategic considerations and trying to exploit the biased bidding behavior.

Even more intriguing, reserve prices appear to react even less sensitively to the precise age of players than the sales prices. The pattern depicted in Figure 3.3a reveals substantial clusters at 0 and at multiples of €50,000. As we discuss in more detail below, this indicates that reserve prices are too clustered as to be compatible with fully rational behavior, where the optimal reserve price is continuous function of the hazard rate of the distribution of buyers' valuations (see e.g. Krishna, 2002). If that was the case, at any age the reserve prices should be similarly dispersed as the final sales prices (see Figure 3.3b).

In the following, we analyze these indications in more detail. After a brief discussion of the estimation model, we present the results from a hedonic regression analysis that allows us to identify the determinants of the reserve price from the set of attributes and the auction details in our data. In addition, we also examine possible interactions between the observed demand and supply side behavior.

3.3.1 Estimation Model and Predictions

If a seller wants to maximize his expected revenue, his choice for the reserve price for a player rationally requires him to form an estimate of the expected bidders valuations. Since both parties, buyers and sellers, share the same information set on the players' attributes when pursuing their evaluation task, our intuition is that similar to the bidders' valuations, also the reserve price is a function of the various player attributes.

To put more structure on the estimation model, consider Figure 3.3a again. Note that the observed discontinuities where the players turn one year older apparently differ in their relative size. To account for this possibility, we decompose the variable *years* into dummies for each age-group, which we label *age17, age18,* and *age19*. For example, the effect of age-group eighteen is measured by *age18* taking value 1 and 0 otherwise. Since these dummies are perfectly correlated, we need to include only two of them in our estimation model. As we drop *age17* from the regression, the resulting coefficients for the included dummies are to be interpreted as the difference in values upon entering an age-group relative to the price of a player aged "17 years 0 days". Moreover, we account for the impact of a marginal day of age separately for each age-group by interacting the age-group dummies with *days*, which yields the variables *days17, days 18,* and *days19* $\in [0, 111]$. They display the days of age conditional on belonging to the specified age-group and zero otherwise. Formally, this corresponds to a piece-wise linear relationship between *totalage* and *askprice*.[20] All other regressors are assumed to enter linearly into the regression model, which is thus given by

$$\begin{aligned} askprice \;=\; \alpha \;&+\; \beta_{age18} \cdot age18 + \beta_{age19} \cdot age19 + \\ &+\; \beta_{day17} \cdot days17 + \beta_{day18} \cdot days18 + \beta_{day19} \cdot days19 + \\ &+\; \beta_{tsi} \cdot tsi + \beta_{acquired} \cdot acquired + \delta \mathbf{X} + u, \end{aligned} \tag{3.1}$$

where $\mathbf{X}$ represents the vector of all other attributes and auction details. With this specification, we are able to identify and measure the average magnitude of potential discontinuities in the reserve price pattern at the players' birthdays. Intuitively, the coefficients for the age-groups (β_{age18} and β_{age19}) reflect the total value difference *across* two subsequent age-groups, while the 112 times the respective impact of a marginal day (β_{day17}, β_{day18}, or β_{day19}, respectively) accounts for the aggregated value decline *within* an age-group. Only if both measures imply the same average decline, the reserve prices evolve continuously in *totalage*. Formally, this corresponds to testable predictions stated in Hypothesis 1.

Hypothesis 1 *If the reserve price pattern exhibits the birthday effect, then*

(i) the coefficient for age18 *is larger than the aggregated value decline per day in age-group seventeen, i.e*

$$|\beta_{age18}| \;>\; 112 \cdot |\beta_{day17}|,$$

(ii) the difference between the coefficients of age19 *and* age18 *is larger than the aggregated value decline per day in age-group eighteen, i.e*

$$|\beta_{age19} - \beta_{age18}| \;>\; 112 \cdot |\beta_{day18}|.$$

[20] This specification is adopted analogously from our analysis for the demand side of the transfer market. For a more detailed discussion see Chapter 2.

The model specification (3.1) also includes the dummy *acquired* to control for possible differences between fresh players and those that have been traded previously. One might be tempted to argue that this should have no effect on the buyers' valuations and therefore should play no role for the sellers' considerations with respect to the reserve price.

However, whether or not a player was sold before conveys a subtle piece of potentially valuable information. To see this, recall from the discussion of *tsi* in Section 3.2 that the skill-levels contain hidden decimal places, while the profile page only shows the adjective corresponding to the integer value for each skill. For example, consider a freshly promoted player (*acquired = 0*) from a manager's youth squad, who displays a *keeping* score of 6. Since skills are completely randomly assigned, his precise skill can take any value within the real interval $[6, 7]$. If we suppose that his sub-skill is uniformly distributed on this interval, a rational buyer would have a prior of 6.5 for the expected *keeping*-skill level.

In contrast to that, if a level-6 keeper was traded previously (*acquired = 1*), there is a positive probability that he was trained in *keeping* and just reached the lower threshold of 6.00 to display a score of 6. Recalling the "train-and-trade" strategy the majority of managers pursues, from perspective of the seller this would be the rational time to offer the player for sale. According to Bayes rule, a rational prospective buyer should adjust his prior of the expected skill-level, and thus his value estimate for such a player, downwards accordingly.[21] Since an optimal reserve price depends on the distribution of the bidder's valuations, if anything, this implies that the reserve prices should not be larger for purchased players relative to fresh ones.

Hypothesis 2 *If the sellers correctly anticipate the considerations of the buyers with respect to a previous sale, the reserve prices for purchased players should not be larger than those for fresh players, i.e.*

$$\beta_{acquired} \leq 0.$$

[21] Since at present about 10% of the *HT* population trains *keeping*, the expected skill for level-6 keeper accordingly reduces to $0.1 \cdot 6.0 + 0.9 \cdot 6.5 = 6.45 < 6.5$

3.3.2 Results

To test the validity of the above hypotheses, we start out with a series of hedonic OLS regressions with the reserve price as the dependent variable, where we only consider observations where the managers set a reserve price different from zero.[22] We either include only the main variables of interest or the full set of controls. In addition, we run a separate regression for experienced managers to see whether they behave differently than the average manager, where we classify a seller as an expert, if his team is ranked among the top 20% in his country (*sellerxp* < 0.2).[23] Moreover, to identify whether the reserve prices follow a similar pattern as the bidders valuations, we additionally run two regressions for the final sales price as the dependent variable, where the first contains all successful trades and the second only those, where the reserve price was set to zero. Table 3.3 presents the results.

Table 3.3: Determinants of Reserve Price and Final Price (OLS)

	Reserve Price (Dep.Var.: *askprice*)			Sales Price (Dep.Var.: *price*> 0)	
	I	II	III (Experts)	IV	V ($askprice = 0$)
days17	-155.3**	-166.94**	-188.39	-1117.71***	-1201.36***
	(70.62)	(70.13)	(179.36)	(65.82)	(151.61)
days18	-101.61***	-99.53***	-161.14	-166.95***	-133.75***
	(37.67)	(37.36)	(104.12)	(25.19)	(41.51)
days19	-37.68	-33.89	-114.01	-30.38	-6.47
	(27.48)	(27.70)	(84.85)	(19.14)	(31.16)
age18	-103154.72***	-105770.16***	-107809.34***	-175188.98***	-174963.87***
	(5648.10)	(5650.93)	(15232.99)	(5114.74)	(11906.80)
age19	-127305.85***	-129555.76***	-134111.3***	-205895.9***	-203476.81***
	(5342.31)	(5466.74)	(14944.07)	(5064.23)	(12023.98)
tsi	31.91***	28.41***	15.59	70.6***	62.83***
	(2.29)	(4.73)	(16.37)	(3.93)	(6.68)
acquired	26590.64***	17617.9***	23255.22**	1549.44	-7125.8
	(1924.51)	(2940.95)	(10404.38)	(2168.67)	(4530.93)
Intercept	107024.18***	98755.29***	105784.1**	122811.82***	146496.82***
	(6537.69)	(14783.43)	(42967.33)	(12407.64)	(19927.25)
skills	no	yes	yes	yes	yes
character	no	yes	yes	yes	yes
daytime	no	yes	yes	yes	yes
weekday	no	yes	yes	yes	yes
R^2	0.47	0.48	0.56	0.74	0.74
N	5108	5108	567	4743	1149
F	453.10	120.72	29.15	244.52	48.03

Notes: Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. "Skills" captures the playing abilities except of *keeping* (= constant). "Character" contains all other player attributes except *tsi*. "Daytime" and "weekday" indicate whether dummies for daytime and day of the week were included.

[22] All results remain qualitatively robust if we instead run a Tobit regression accounting for our sample being left-censored at zero. Moreover, the same holds true if we consider log-linearized reserve records and control for influential outliers using a robust regression procedure. For the sake of brevity, here we omit the regression tables but they are available from the authors upon request.

[23] Using *sellerxp* < 0.2 leaves us with a bit less than 30% of the observations from the subsample where we were able to construct the experience dummy.

The birthday-effect. In any reserve price regression (columns I-III), the age of a player has an highly significant negative impact on the level of the minimum bid set by the sellers. Representatively focus on the full-control specification in column II. The coefficient of *age18* indicates that the average reserve price for a player who just turned eighteen is substantially lower (by €105,770) than that for a player aged "17 years 0 days" at 99%-significant t-statistics, which is not too surprising given the game's training algorithm. However, holding all other variables constant, a marginal day within the age-group seventeen just accounts for a decline of €167 in the average reserve price. Aggregated over the whole year (112 days), this gradual day-by-day decline explains only 18% (€18,704) of the total value loss measured on the eighteenth birthday of a player. More precisely, the remaining 82% of the total decline establish an enormous discontinuity in the reserve price pattern where a player turn eighteen, i.e. $|\beta_{age18}| \gg 112 \cdot |\beta_{days17}|$. A Wald-test shows that this difference is significant on the highest level (p-value: 0.000). Similarly, also at the nineteenth birthday of a player we find a significant discontinuity of 53% of the total decline between the age-groups eighteen and nineteen, and thus $|\beta_{age19} - \beta_{age18}| \gg 112 \cdot |\beta_{days18}|$.

In the regression for experienced sellers (column III), observe that the impact of a marginal day is qualitatively similar but no longer significant, which may be due to the reduced number of observations (N=567). Note that this implies that the reserve prices do not adjust to age *within* but only *across* age-groups, making our result even stronger. If we account for the aggregate day effects despite their insignificance, we find a discontinuity of 80% and 31% at the eighteenth and nineteenth birthday, respectively. The results from the regression analysis are thus consistent with the prediction from Hypothesis 1 that reserve price pattern picks up the birthday effect, which establishes our first result.

Result 1 (The birthday effect): *Similar to the sales prices, the reserve price pattern does not evolve continuously in the total age of a player. Instead, it picks up the birthday effect in form of substantial and highly significant discontinuities at the players' birthdays.*

Reserve price and bidder valuation. Observe further that adding the full set of controls to the regression (column II) only slightly improves the predictive power of the estimation model relative to column I, implying that the age variables, *tsi*, and *acquired* are indeed the most influential factors for the choice of the reserve price.[24] Moreover, except for *acquired*, all coefficients in the reserve price regressions qualitatively mirror those for the sales prices in columns IV and V.[25] This is in line with the prediction that sellers take into account the bidders' valuations when forming their reserve price.

To distinguish whether the supply side reacts to the demand side behavior or whether the causality is reversed, in column V we only consider the sales prices from auctions where the seller set no reserve price ($askprice = 0$). If the birthday effect persists also for this subsample, we can rule out that the latter originates from the supply side of the transfer market. As it turns out, this is indeed the case. We find clear evidence for a birthday effect in the sales prices in column V, amounting to highly significant discontinuities of 23% and 47% of the total decline at the eighteenth and nineteenth birthday, respectively.[26] Since the bidding pattern is thus qualitatively unaltered in absence of a positive reserve price, this implies that the sellers take into account the expected bidders' valuations when making their reserve price choice, yielding our second result.

Result 2 (Reserve prices relate to bidder valuations): *The reserve prices are shaped remarkably similar to the final prices and share the same subset of influential player attributes. The finding of a birthday effect for the supply side is consistent with the sellers relating their choice of the reserve price to the bidders' valuations.*

Reserve price and auction outcome. Before we go on with our analysis of seller behavior, it proves useful to briefly consider the predicted effects of a reserve price on the final price on a more general level. Relative to a zero minimum bid, one of the most basic general predictions is that higher reserve prices should reduce the likelihood of a successful sale, because low-value bidders will cease to participate (Reiley, 2006). Since all but one of the 24% failed auctions in our sample exhibited a strictly positive minimum bid, this prediction is clearly met.[27] A non-parametric Wilcoxon-Mann-Whitney (WMW) ranksum

[24]While each group of controls turns out to be jointly significant, all effects of individually significant control variables are of secondary order and do not conflict with any of our results. To ease the exposition, we therefore omit a detailed discussion. The full regression tables can be requested from the authors.

[25]Note however that the relevant coefficients in the price regressions are quantitatively much larger as in those for the reserve price and the latter does not react as nuancedly to the precise age. As we will discuss below, this likely due to the strong clustering of the reserve prices.

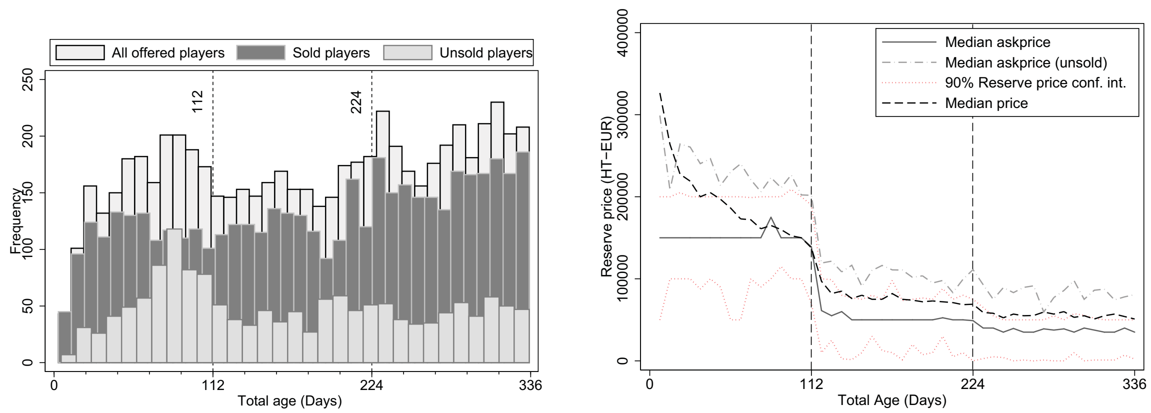
[26]The discontinuities for the full sample of non-zero sales prices can be readily calculated from column IV in Table 3.3 and are given by 29% ($\frac{\beta_{age18}-112\cdot\beta_{day17}}{\beta_{age18}}$) and 39% ($\frac{\beta_{age19}-\beta_{age18}-112\cdot\beta day18}{\beta_{age19}-\beta_{age18}}$), respectively.

[27]Since the sample was hand-collected, this one observation may be due to a data-entry error.

test confirms that the askprices for unsold players are significantly higher than those for successful trades (p-value: 0.000). We also find that the expected final price unconditional on a successful sale is lower if the reserve price is above zero (p-value: 0.000).

Evidence for sophisticated seller behavior. Conditional on a successful sale, however, the expected price is larger for positive askprices (p-value: 0.000). Intuitively, if the seller manages to set a reserve price between the highest bidder's and the second highest bidder's valuation, the auction price is mechanically higher than if there was no minimum bid and the seller successfully reaps some of the winner's surplus.[28] Since in 15.9% of the successful trades there was only a single bidder and the winning bid equaled the reserve price (Panel B of Table 3.2), we take it that the sellers in these 756 auctions were successful in their attempt to appropriate some of the highest bidder's surplus.

Figure 3.4: Distribution of Sold and Unsold Players and Median Reserve Prices



(a) Sold and Unsold Players by Age

(b) Mean Reserve Price by Age

Notes: The dashed lines at 112 and 224 mark the points where the players turn eighteen and nineteen, respectively.

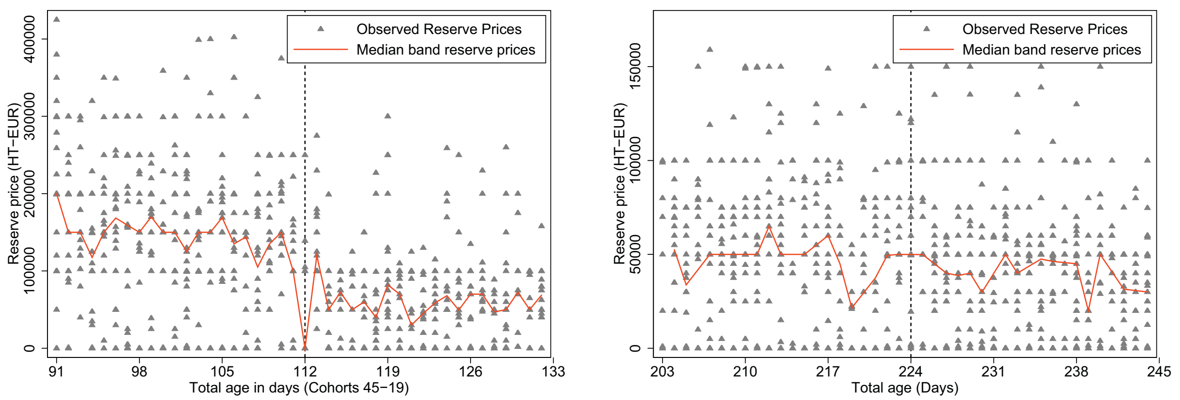
A natural next step is to ask whether at least some sellers react strategically in anticipation of the biased bidding behavior on the demand side of the transfer market. We would expect that sellers who are aware of the "birthday effect" should rationally try to sell players that are close to turn one year older, thereby avoiding to bear the accompanying value loss themselves. Consistent with this intuition, the distribution for sold and unsold players shown in Figure 3.4a clearly indicates an increased number of sale offers shortly before players turn eighteen. At about 90 days of total age, both the total number of offers and the number of failed

[28]See also Trautmann and Traxler (2008), who distinguish between this mechanical effect of "surplus appropriation" and a potential psychological channel of influence of the reserve price on the final price, as suggested in Rosenkranz and Schmitz (2007).

auctions increase substantially, while the number of sales (and also the selling price pattern) remains largely constant. At the same time, the median askprice exhibits a local peak at exactly the same age level (Figure 3.4b). Though considerably less pronounced, we observe an similar increase around three weeks before the nineteenth and also the twentieth birthday.

Our intuition is that the sellers set rather high reserve prices at that stage in the hope to find a buyer who is not aware of the birthday effect and is thus willing to pay the asked price. Clearly, the downside of this strategy is that the probability for a successful sale is considerably reduced. This would also explain the higher number of failed auctions we observe. In even closer proximity to a player's birthday, however, a seller should rationally change his strategy and try to maximize the probability for a successful sale by charging a rather low minimum bid.

Figure 3.5: Reserve Prices in Close Proximity to Discontinuity Point



(a) Eighteenth Birthday (at 112)   (b) Nineteenth Birthday (at 224)

Notes: The figure includes 756 (883) observations of players 3 weeks before and after their eighteenth (nineteenth) birthday.

Figure 3.5 plots the reserve prices closely before and after the eighteenth (nineteenth) birthday of the players in our sample. Consistent with our intuition, observe that the median reserve price for seventeen year old players begins to decline at 110 days of total age and drops to zero at 112, while it moves upwards again immediately after the birthday. A similar effect exists for the nineteenth birthday, though the drop is not as emphasized and takes place a few days before the birthday.[29] Though merely inferred by inspection, we take these observations as an indication that at least some sellers behave strategically.

[29]However, note that the absolute impact of the birthday effect is considerably lower at age nineteen, which might explain the reduced reaction of the sellers at this point.

Result 3 (Strategic seller behavior): *At least some sellers show quite sophisticated strategic considerations in their choice of the reserve price, as indicated by an increased number of sale offers and low minimum bids in close proximity to the birthdays. Moreover, a substantial fraction of sellers manages to extract additional rents from the winning bidder by setting a reserve price above the second highest bidder's valuation.*

The impact of a previous sale. Next, we consider the impact of a previous sale of a player on the reserve price. Returning to Table 3.3, we find strong evidence that sellers set significantly higher minimum bids for players they acquired on the market ($acquired = 1$) relative to fresh players from their own youth squads ($acquired = 0$). Throughout all reserve price specifications (columns I-III), contrasting with Hypothesis 2 the coefficient $\beta_{acquired}$ is large and significant on the highest level, implying a strong positive correlation of *acquired* with *askprice*. Holding all other variables constant, the average reserve price for purchased players in column II exceeds that for fresh players by an amount of €17,618, or 23% of the mean reserve price. To illustrate the economic significance of this effect, note that the statistically weighted average impact of *tsi*, which we obtain by multiplying its coefficient ($\beta_{tsi} = 28.4$) times one standard deviation of the average *tsi*-score ($\sigma_{tsi} = 428$), only amounts to €12,155, i.e. about two-thirds of the coefficient for the dummy *acquired*.

In contrast to that, the regression for *price* conditional on sale in (column IV) yields an insignificant coefficient for *acquired*, all else equal implying that the buyers' valuations do not significantly differ whether or not a seller was traded before. Furthermore, in the specification for auction prices with a zero reserve price (column V), the coefficient for *acquired* is large and has a negative sign, i.e. qualitatively the effect goes into the opposite direction for buyers. Though not statistically significant, this points towards the latter on average having a higher willingness to pay for fresh players, which is consistent with our above reasoning that a previous sale rationally translates into a lower expected (sub-) skill-level.[30]

An immediate implication of these results is that auctions for purchased players are ceteris paribus more likely to fail. Among the players that were bought ($acquired = 1$), the share of failed auctions was 32.6%, which is substantially higher than that for fresh players at 20.3%.

[30]Table 3.7 in Appendix 3.6 shows the results from an additional Tobit regression, where we include also the failed auctions with a zero sales price, i.e. with the sales price unconditional on entry as the dependent variable. In this approach, *acquired* has a highly significant negative coefficient ($\beta^{tobit}_{acquired} = -13,579$) which accounts for roughly 17% of the unconditional mean price. Unconditional on a successful sale, the average sales price is thus considerably lower for purchased players.

A Pearson's chi-square test confirms that there is a statistically significant relationship between *acquired* and the frequency of players remaining unsold (p-value: 0.000). Moreover, if a player was previously traded, we find that in only 9.9% of the cases the reserve price was set to zero. Among fresh players with 22.4% this share is more than twice as large. To further substantiate this finding, we estimate the likelihood for a positive reserve price in a series of Logit regressions on the dummy *d_ask*, which takes the values 1 or 0 depending on whether or not there was a positive minimum bid. The results are shown in Table 3.4.

Table 3.4: Likelihood of Non-Zero Reserve Price

	Logit I		Logit II		Logit III expert sellers	
Dep.Var.: *d_ask*	Odds Rt.	Marg. Efct.	Odds Rt.	Marg. Efct.	Odds Rt.	Marg. Efct.
days17	1.0032	0.0005	1.0029	0.0004	0.9969	-0.0005
	(0.0024)		(0.0024)		(0.0059)	
days18	0.9980	-0.0003	0.9976	-0.0003	0.9991	-0.0001
	(0.0018)		(0.0018)		(0.0047)	
days19	0.9983	-0.0002	0.9980	-0.0003	0.9978	-0.0004
	(0.0015)		(0.0015)		(0.0044)	
age18	0.8622	-0.0211	0.8516	-0.0223	0.4291	-0.1368
	(0.1726)		(0.1727)		(0.2232)	
age19	0.7563	-0.0397	0.7673	-0.0368	0.4271	-0.1376
	(0.1434)		(0.1495)		(0.2231)	
tsi	0.9999*	0.0000	0.9998	0.0000	1.0007	0.0001
	(0.0001)		(0.0002)		(0.0006)	
acquired	2.4449***	0.1270	1.8536***	0.0857	4.7540***	0.2521
	(0.2101)		(0.227)		(2.2669)	
skills	no		yes		yes	
character	no		yes		yes	
daytime	no		yes		yes	
weekday	no		yes		yes	
LR Chi	212.93		306.42		95.04	
N	6258		6258		748	

Notes: The Logit procedure estimates the impact of the independent variables on the probability to observe "*d_ask* = 1" relative to "*d_ask* = 0". For each specification, the left (right) column states the odds ratio (marginal effect) for the respective regressor variable. Standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. Sellers are classified as experts if *sellerxp*<0.2.

The first two columns (Logit I) show the odds ratios (exponentiated coefficients) and the corresponding marginal effects (instantaneous change) in the probability when only the main variables are used as regressors. Note that *acquired* is the only variable which has an highly significant impact. For a purchased player, the odds for a strictly positive reserve price (versus a zero reserve price) increases by a highly significant factor of 2.4 as compared to a player that was internally promoted from a seller's own youth squad. In terms of the marginal effect, a non-zero reserve price is 12.7% more likely for previously acquired players.[31] If we employ the full vector of player attributes and auction details as controls (Logit II), all

[31]The marginal change equals the partial derivative of the predicted probability with respect to the respective regressor variable.

results remain qualitatively robust. In line with our previous results, the same holds true in the separate regression for expert sellers (Logit III). If anything, the effect seems to be even more pronounced for experts.

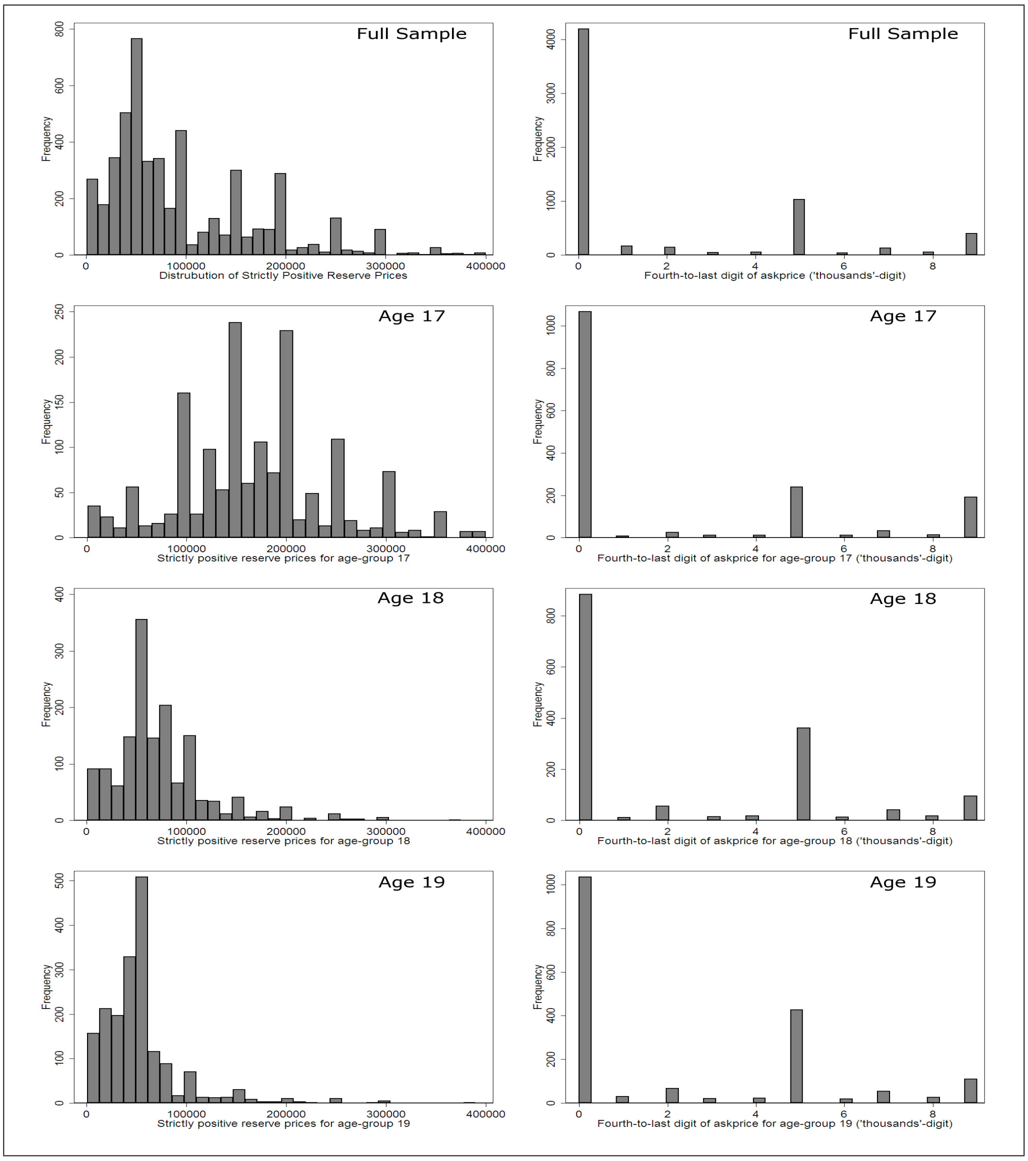
Hence, if a player was bought rather than promoted internally by the seller, not only the level of the reserve price is higher on average, but also the likelihood that it is set different from zero at all. Stated differently, even though the managers interact in a highly competitive market environment, our findings indicate that the sellers in *HT* exhibit some form of an entitlement effect with respect to players that they acquired on the market, but not for those promoted from their own youth squad.

Result 4 (Entitlement effect for acquired players): *Relative to internally promoted players, the average seller in* HT *demands a positive reserve price premium for players they previously acquired on the market* ($\beta_{acquired} > 0$). *Hypothesis 2 can thus be rejected.*

This finding suggests that the sellers of purchased players fall prey to the *sunk cost fallacy* and *loss aversion*. For instance, a seller might be tempted to regard his own acquisition cost as a benchmark for the reserve price he sets and thus charges at least the same amount, or feels even entitled to demand an additional premium. Yet, this cost is sunk and should rationally not affect his choice. Moreover, since his valuation was the highest in the previous auction, this might be too high a threshold, if a player hasn't significantly increased in quality in the meantime. In contrast, for internally promoted players there is no such benchmark, which would explain the difference in reserve prices that we observe.

Reserve price clustering. While we find final and reserve prices to respond very differently if a player has been traded previously, we already pointed out that the impact of all other influential variables is *qualitatively similar* for both. However, all coefficients for the age variables and also for *tsi* are *quantitatively much larger* in the price regressions than in those for the reserve price. For instance, in column IV of Table 3.3, holding all other variables constant, a marginal day in age-group seventeen reduces the sales price on average by €1,118, while the analogous effect for the reserve price is only €167, or 15% of the former. Intuitively, the minimum bids react substantially less sensitively to the precise age. A possible explanation for this pattern arises from the fact that the reserve prices are remarkably clustered around focal points at multiples of €50,000, as indicated in Figure 3.3a above. This intuition is further strengthened by an inspection of the frequency distribution of (non-zero) reserve prices in our sample as depicted in Figure 3.6.

Figure 3.6: Distribution of Reserve Prices and their Fourth-to-last Digits



First, consider the four histograms on the left of Figure 3.6 which depicts the frequency distribution of (non-zero) reserve prices for the full sample and for each age-group separately. Though they naturally exhibit some variation, we find that reserve prices are *substantially clustered* at multiples of €50,000. The distribution for the full sample exhibits several distinct spikes ranging from €50,000 to €300,000. A similar pattern arises on the individual age-group levels, where the effect is most accentuated for seventeen year old players. In particular, their reserve prices are most frequently set to the (round) values of €100,000, €150,000, and €200,000. The patterns for ages eighteen and nineteen also exhibit strong focal points at €50,000 and at some of its multiples, though less pronounced due to the age-induced decline in the players' values.

Second, we also find evidence for considerable *lower scale clustering*. On the right of Figure 3.6 we plot the frequency distribution of the fourth-to-last digit (the "thousands"-digit) of the reserve prices. For example, if the reserve price is given by €6**7**,000 the digit takes value "7". Accordingly, the spikes at the values of 0, 5, and 9 indicate that a majority of reserve prices were set at multiples of €5,000, or just below the next full ten-thousand.

Result 5 (Clustering at round numbers): *The distribution of reserve prices exhibits substantial clusters at multiples of €50,000, and also on a lower scale at multiples of €5,000.*

According to Sonnemans (2006), such (large scale) round number-clustering could be caused by boundedly rational sellers, who form mental target prices, which then serve as a "good enough"-solution in their view instead of considering the precise distribution of bidders' valuations. In line with this intuition, and also with the findings documented in Benartzi and Thaler (2008) for the determination of savings choices, we interpret the strong clustering in the reserve price pattern as evidence for sellers using a round number heuristic, or rule-of-thumb, when making their choice for a reserve price to considerably simplify their decision making.

Moreover, recall that a large fraction of the sellers (18%) sets a reserve price of zero. Both the strong clustering and frequent absence of a positive minimum bid stand in marked contrast to the theoretical predictions for the optimal reserve price in an IPV context. This suggests that many sellers do not efficiently utilize the reserve price as an instrument to maximize their expected revenue. Intuitively, an efficient strategy to appropriate a share of the winning bidder's rent would require a seller to form very fine-tuned estimates of the expected (second) highest-order statistic derived from the distribution of valuations rather than to employ a simplifying rule-of-thumb. Moreover, due to the large scale clustering a substantial fraction of reserve prices will be set considerably below and above the optimal level. In the latter case, a more than optimal number of auctions will fail resulting in an inefficient allocation.

3.4 Suboptimal Reserve Price and Foregone Revenue

In this section we quantify the economic consequences of suboptimal reserve prices by providing estimates of how much expected revenue is effectively lost as compared to the situation with optimal reserve prices. In doing so, we account for the effect of age on the players' values, and thus on the level of the optimal reserve price, by subdividing our sample into weekly intervals of precise age, yielding a total of 48 "cohorts". Each age-group consists of 16 cohorts, where the intervals are defined such that they only contain players from the same age-group.[32] Abstracting from the small variation of the precise age, we assume that the values for the players within cohort j are realizations from the same underlying value distribution function, $F_j(v)$. Under this assumption, the optimal reserve price, r_j^*, is the same for all players belonging to cohort j.

In the following, we start with a brief description of our approach to determine the optimal reserve price.[33] In particular, we first use the information on the observed sales prices in our data to identify a non-parametric estimate for the valuation primitive of each cohort. By using a Maximum Likelihood method (MLE) to obtain a parametric distribution best describing these point estimates, we then determine the optimal reserve prices and the maximum expected revenues. Subsequently, we present the results from comparing the expected revenues at the actual reserve prices to the corresponding optimal benchmark.

3.4.1 Estimation of the Optimal Reserve Price

As a first step to estimate the optimal reserve price it proves useful to re-consider the seller's problem in general. In particular, the expected revenue of a seller who sets a reserve price of r for an object for sale is given by

$$\Pi(r) = v_0 F(r)^N + rN(1 - F(r))F(r)^{N-1} + \int_r^{\overline{v}} u(1 - F(u))(N - 1)F(r)^{N-2} f(u)du, \qquad (3.2)$$

[32]For example, cohort 16 includes all players in the interval *totalage* $\in [105, 111]$ (i.e. "17 years 105 days" to "17 years 111 days"), while cohort 17 contains those with *totalage* $\in [112, 118]$ (i.e. "18 years 0 days" to "18 years 6 days"). Hence, the birthday effect will take place across but not within the cohorts.

[33]For the sake of brevity, we omit the details of the calculus, which was performed with Matlab. The corresponding m-files can be requested from the authors. All proofs for the validity of the applied approach are available in standard textbooks, e.g. Paarsch and Hong (2006) and Krishna (2002).

where $0 \leq v_0 \leq r$ is the reservation value of the seller and N the number of potential bidders. Under the independent private value (IPV) assumptions, the bidders' individual valuations are i.i.d. draws from the increasing distribution function $F(v)$, the "valuation primitive", which has support $[0, \overline{v}]$ and admits a continuous density function $f(v) \equiv F'(v)$ (see e.g. Krishna, 2002). The first term on the RHS of (3.2) thus captures the expected utility from the event that none of the N potential buyers realizes a valuation above r and the seller gets v_0, which occurs with probability $F(r)^N$. The second term reflects the event where all but one bidder have a lower valuation than r, which occurs with probability $N(1 - F_j(r))F_j(r)^{N-1}$. In that case, the winning bid equals r and the seller successfully manages to reap some of the winner's surplus. Finally, the third term states the expected sales price conditional on more than one bidder having a valuation of at least r. The choice for the reserve price thus involves a trade off between a higher probability that the auction fails and realizing additional gains from a sales price above the second-highest bidder's valuation.

As a general result under IPV, for any arbitrary value distribution function $F(v)$, the optimal reserve price r^* that maximizes (3.2) is independent of the number of potential bidders and given by

$$r^* = v_0 + \frac{1}{\mu(r^*)}, \tag{3.3}$$

where $\mu(v) \equiv \frac{1-F(v)}{f(v)}$ is the hazard rate of $F(v)$ (see Riley and Samuleson, 1981). Thus, the optimal reserve price r^* will always lie above v_0 and depends on the valuation primitive of the bidders, which is unknown the econometrician.[34]

However, $F(v)$ is non-parametrically identified from the winning bids, i.e. the observed sales prices (Athey and Haile, 2002). Depending on the auction format, the latter describe the empirical distribution of the i^{th} order statistic from an i.i.d. sample of size N from the valuation primitive $F(v)$. Since the auction format in *HT* is strategically equivalent to a sealed-bid-second-price auction, the winning bid equals the second-highest bid plus one discrete increment, whenever the reserve price was set below the second highest bidder's valuation.[35]

[34] In Chapter 2 we argue that the managers do not necessarily know their true valuation, v. Instead, they base their bids on the *expected* valuation for a player, $E(v)$, which results from their individual evaluation of his attributes. Strictly speaking, the optimal reserve price thus depends on the primitive of the *expected* valuations, $F[E(v)]$. To simplify the notation, we continue to use the term $F(v)$.

[35] In a sealed-bid second price auction, the equilibrium bidding strategy is to place a bid equal to one's own valuation, v_i, while in the English ascending open bid auction a bidder will repeatedly increase his bid until the current price reaches v_i and exit thereafter.

Denote by w_{jk} the observed sales price of auction $k = 1, ..., K_j$ in cohort j. Hence, the K_j sales prices observed for cohort j describe an empirical distribution function $\hat{F}_j(w_{jk})$,[36] which is equivalent to the distribution of the second-highest order statistic of the valuation primitive, i.e. the distribution of the second-highest valuations $F_{(2)j}(v, N_{jk})$:

$$\hat{F}_j(w_{jk}) \equiv F_{(2)j}(v; N_{jk}) = N_{jk} \cdot F_j(v)^{N_{jk}-1} - (N_{jk} - 1) \cdot F_j(v)^{N_{jk}}, \tag{3.4}$$

where N_{jk} is the number of potential bidders in auction k in cohort j. Though the reserve price is independent of N_{jk}, it hence enters the distribution of the second-highest order statistic, which in the following is used to identify the valuation primitive $F_j(v)$. However, we have no information on how many potential bidders will view a particular player. At this point, we therefore need to make a simplifying assumption on the value of N_{jk} used to obtain the estimate of $F_j(v)$.

Assumption 1 (Potential number of bidders): *The number of potential bidders is exogenously fixed and identical for all auctions in the sample:* $N_{jk} = N \; \forall \; j, k$.

In addition, we borrow on the dataset of 17,510 HT-auctions from the study of the demand side in Chapter 2 to determine appropriate values for N to be used in the estimation. Conveniently, these data contain information on the number of bids placed in each auction.[37] We argue that this is a reasonable proxy for the potential number of bidders viewing an individual player on the transfer market. In particular, we obtain the point estimates of $F_j(v)$ for $N = 7$, which is equal to the median number of bids in these data.[38]

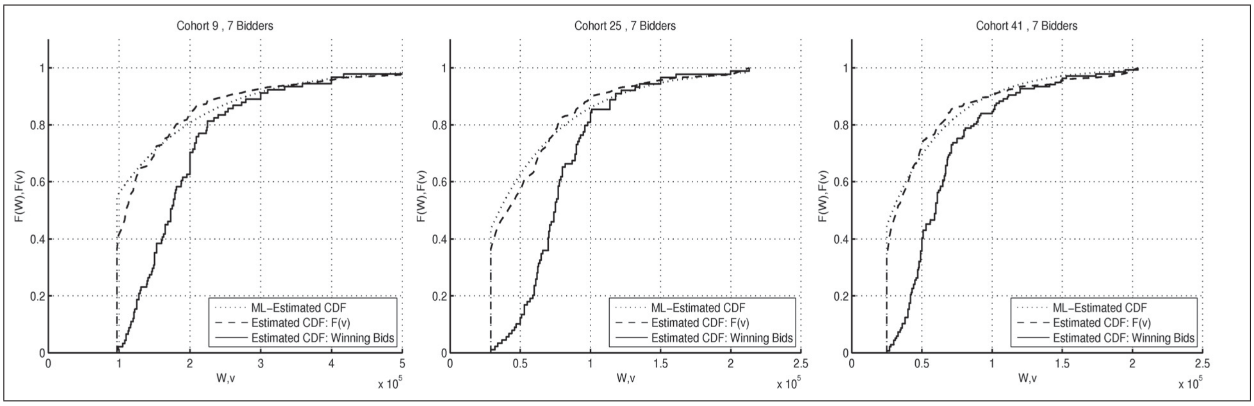
By substituting $N = 7$ and the values for $\hat{F}_j(w_{jk})$ into (3.4), we are able to obtain point estimates of the valuation primitive for each of the 48 cohorts by numerically solving for the roots with respect to $F_j(v)$. As the results are qualitatively similar for all j, Figure 3.7 shows the results from this approach for a representative cohort in the middle of each age-group.

[36] Given a vector of length T, an empirical distribution function is a cumulative probability distribution function that concentrates a probability of $1/T$ at each of the T elements in the vector.

[37] For a graphical illustration of the distribution of the number of bids see Figure 3.10 in Appendix 3.6.

[38] To account for the effect of the number of potential bidders we repeated the estimation for $N = 13$ (mean number of bids), $N = 2$ (25th-percentile), and $N = 18$ (75th-percentile). For a discussion of the results see Section 3.4.2 below. All tables are available from the authors upon request.

Figure 3.7: Distribution Functions of First- and Second-Highest Order Statistic and ML-Estimator



The figure depicts the empirical distribution $\hat{F}_j(w_{jk})$ of the winning bids and the corresponding point estimates of the valuation primitive $F_j(v)$ for cohorts $j = \{9, 25, 41\}$. For every j, $\hat{F}_j(w_{jk})$ resembles the cumulative distribution function (CDF) of a normal distribution, and $F_j(v)$ is approximately exponentially distributed. Assuming the valuation primitive indeed follows an exponential distribution, we have that

$$F_j(v) \xrightarrow{\text{MLE}} \tilde{F}_j(v; \lambda_j) = 1 - e^{-\lambda_j v} \quad \forall \;\; v \in \mathbb{R}_0^+ \tag{3.5}$$

where λ_j is the Maximum-Likelihood estimator (MLE) for the rate parameter of the exponential distribution $\tilde{F}_j(v; \lambda_j)$ best describing $F_j(v)$, which is also depicted in Figure 3.7.[39]

Having derived an estimator for the underlying valuation primitive $F_j(v)$ for the players in cohort j, we are able to determine the corresponding optimal reserve price r_j^* up to a constant, i.e. the reservation value v_0 of the seller. However, v_0 is not observed and thus unknown. Therefore, the estimation results for the optimal reserve prices additionally rely on the following simplifying assumption.

Assumption 2 (Reservation Values): *The reservation value v_0 is constant across all sellers and given by the consumption value (CV) of the players in the sample.*

In particular, the fact that a player is offered for sale implicitly signals that his current owner does not consider him as a suitable trainee for his current training-strategy. Given this rationale, our belief is that it is plausible to assume that the value he attaches to the player is likely to reflect the latter's CV. As a proxy for the CV of the players in our sample,

[39]Given the exponential formulation and k observations in cohort j, the log-likelihood function used to estimate λ_j depends on the density of the second-highest-order statistic, $\tilde{f}_{(2)j}(v, \lambda_j)$, and is given by $\ell(\lambda_j) = \sum_k ln(\tilde{f}_{(2)j}(v, \lambda_j)) = \sum_k \left[ln\big(N(N-1)\big) + ln\big(1 - \tilde{F}_j(v, \lambda_j)\big) + (N-2)ln\big(\tilde{F}_j(v, \lambda_j)\big) + ln\big(\tilde{f}_j(v, \lambda_j)\big)\right]$.

we take about 80% of the average sales price (€ 48,921) observed for the oldest players, i.e. cohort 48, yielding an assumed value of $v_0 = 40,000$ for the reservation value of the sellers.[40]

Table 3.5 shows the resulting estimates of the optimal reserve price for each cohort, which are obtained by substituting v_0 and $\tilde{F}_j(v;\lambda_j)$ into (3.3). In addition, we state the number of observed winning bids, on which the identification of $\tilde{F}_j(v;\lambda_j)$ is based. Except for cohorts $j=1$ and $j=2$, the youngest players in the sample, all estimates are based on more than 70 observations, plausibly yielding considerably precise descriptions of the underlying valuation primitives.

Table 3.5: Optimal and Actual Reserve Prices by Cohort

Age 17					Age 18					Age 19				
j	K_j	r_j^*	r_j^{med}	r_j^{mean}	j	K_j	r_j^*	r_j^{med}	r_j^{mean}	j	K_j	r_j^*	r_j^{med}	r_j^{mean}
1	23	241,112	150,000	135,000	17	73	109,036	61,250	68,395	33	124	82,590	40,000	45,727
2	26	229,899	150,000	151,845	18	92	96,575	55,000	66,860	34	131	80,877	40,000	39,464
3	73	198,265	150,000	154,250	19	84	96,970	60,000	60,121	35	122	77,021	35,000	39,257
4	102	189,391	150,000	150,026	20	99	92,687	50,000	50,188	36	119	80,142	39,500	38,888
5	90	178,115	150,000	146,889	21	93	96,593	50,000	50,333	37	105	77,077	35,000	38,167
6	76	177,689	150,000	152,053	22	91	92,133	50,000	60,841	38	100	78,665	35,000	36,332
7	102	171,840	150,000	138,127	23	102	91,521	50,000	52,082	39	112	81,601	39,000	34,705
8	108	165,431	150,000	141,765	24	109	101,381	50,000	62,222	40	110	79,073	37,500	37,782
9	91	161,678	150,000	145,988	25	89	90,902	50,000	52,565	41	137	82,155	39,001	39,956
10	105	158,699	150,000	145,404	26	95	93,371	50,000	55,163	42	120	78,161	35,000	43,281
11	80	152,832	150,000	145,500	27	77	88,999	50,000	54,450	43	113	79,655	40,000	38,767
12	87	153,830	175,000	162,195	28	73	88,684	50,000	53,206	44	132	76,097	37,500	39,126
13	87	148,599	150,000	159,118	29	83	91,992	52,500	58,544	45	146	78,623	35,000	41,171
14	84	151,867	150,000	157,731	30	105	88,497	50,000	47,912	46	139	78,166	35,000	34,869
15	87	147,818	150,000	145,702	31	106	88,257	50,000	56,641	47	109	76,930	40,000	37,731
16	87	133,922	139,001	130,970	32	101	85,944	49,000	50,404	48	137	78,602	35,000	37,961

As we would expect, an inspection of the stated values for r_j^* shows a clear tendency of the level of the optimal reserve price to decline as age increases. Note that the estimates for $F_j(v)$, from which the respective values for r_j^* are calculated, are solely based on the observed sales prices in cohort j, which may be subject to considerable variation. Considering this fact, it is not surprising that the decline is not strictly monotonic across all cohorts. Importantly, however, observe that the estimated optimal reserve price exceeds the median (r_j^{med}) and mean (r_j^{mean}) of the observed minimum bids for almost all cohorts, and often quite substantially.

[40] While the age-dependent APV will be very small for players at the end of age-group nineteen, in Chapter 2 we still find evidence for a small birthday effect at the age of twenty. Therefore, we adjust our estimate of the CV downward to the value of € 40,000. All results remain robust, if we instead employ the average sales price of the oldest players in age-group nineteen as a proxy for v_0.

3.4.2 Expected Revenue at Optimal and Actual Reserve Prices

In Section 3.3 we have shown that a substantial fraction of sellers sets either no reserve price or uses a round-number heuristic to simplify their decision making. For example, note that 22 (17%) of the 130 sellers in cohort $j = 9$ set a reserve price of zero, 17 (13%) opted for a value of €150,000, 11 (8%) chose €200,000, and the values of €100,000 and €250,000 were each observed in 6 (5%) auctions, all together accounting for about half of the sellers. The patterns for the other cohorts are quite similar. Therefore, we are particularly interested how the expected revenues for these sellers compare to the estimated optimum.

Given the approximated valuation primitive $\tilde{F}_j(v;\lambda_j)$, analogously to (3.2) the expected revenue at a reserve price of r_j for a player from cohort j is given by

$$\tilde{\Pi}(r_j) = v_0\tilde{F}_j(r)^N + rN(1-\tilde{F}_j(r))\tilde{F}_j(r)^{N-1} + \int_{r_j}^{\overline{v}} u(1-\tilde{F}_j(u))(N-1)\tilde{F}_j(u)^{N-2}\tilde{f}_j(u)du, \quad (3.6)$$

where we denote $\tilde{F}_j(v) \equiv \tilde{F}_j(v;\lambda_j)$ to simplify the notation. For the values of N and v_0 assumed above, the maximum expected revenue for each cohort is given by substituting r_j^* into (3.6), where we additionally employ the 99th-percentile of the observed sales price in each cohort as a proxy for $\overline{v}$. Similarly, we are able to determine the expected revenue for any reserve price observed in our sample. By comparing the resulting outcomes to the respective optimum, we are thus able to determine the shares of expected revenue lost due to deviations from the optimal reserve price.

Table 3.6 states results for age-group seventeen at several preeminent points of the actual reserve price pattern, including zero and the main cluster points as indicated in Figure 3.6.[41] To simplify the notation, we refer to a cluster point by C_x, where x indicates its respective magnitude.

At a first glance, the deviations from the optimum appear to be small. The expected revenues at the mean and median of the actual reserve prices (columns II and III) are remarkably close to $\tilde{\Pi}(r_j^*)$. If we consider the sellers who abstain from setting a positive reserve price (column IV), we find that they only lose a share of about 2% relative to the optimum. Moreover, also the deviations at the cluster points C_{50} and C_{100} (columns V and VI) are of similar

[41] The corresponding results for age-groups eighteen and nineteen are qualitatively similar as shown in Tables 3.8 and 3.9 in Appendix 3.6.

Table 3.6: Expected Revenues for Age-Group 17 (Cohorts 1-16)

		I	II	III	IV	V	VI	VII	VIII
Coh.	Obs.	$\tilde{\Pi}(r_j^*)$	$\tilde{\Pi}(r_j^{mean})$	$\tilde{\Pi}(r_j^{med})$	$\tilde{\Pi}(0)$	$\tilde{\Pi}(C_{50})$	$\tilde{\Pi}(C_{100})$	$\tilde{\Pi}(C_{200})$	$\tilde{\Pi}(C_{250})$
1	24	263,433	260,425	260,858	259,570	259,575	259,791	262,635	263,383
			(-1.1%)	(-1.0%)	(-1.5%)	(-1.5%)	(-1.4%)	(-0.3%)	(-0.0%)
2	29	264,041	261,800	261,734	260,208	260,215	260,488	263,552	263,747
			(-0.8%)	(-0.9%)	(-1.5%)	(-1.4%)	(-1.3%)	(-0.2%)	(-0.1%)
3	89	236,296	235,121	234,937	232,506	232,523	233,071	236,293	233,470
			(-0.5%)	(-0.6%)	(-1.6%)	(-1.6%)	(-1.4%)	(-0.0%)	(-1.2%)
4	127	225,672	224,625	224,623	221,878	221,900	222,573	225,563	221,372
			(-0.5%)	(-0.5%)	(-1.7%)	(-1.7%)	(-1.4%)	(-0.0%)	(-1.9%)
5	104	213,323	212,543	212,673	209,507	209,539	210,419	212,779	206,451
			(-0.4%)	(-0.3%)	(-1.8%)	(-1.8%)	(-1.4%)	(-0.3%)	(-3.2%)
6	97	197,134	196,581	196,499	193,318	193,350	194,239	196,565	190,148
			(-0.3%)	(-0.3%)	(-1.9%)	(-1.9%)	(-1.5%)	(-0.3%)	(-3.5%)
7	131	207,082	206,140	206,643	203,246	203,286	204,309	206,095	198,386
			(-0.5%)	(-0.2%)	(-1.9%)	(-1.8%)	(-1.3%)	(-0.5%)	(-4.2%)
8	139	194,449	193,908	194,203	190,583	190,634	191,829	192,816	183,568
			(-0.3%)	(-0.1%)	(-2.0%)	(-2.0%)	(-1.3%)	(-0.8%)	(-5.6%)
9	130	195,223	194,958	195,072	191,336	191,394	192,704	193,105	182,905
			(-0.1%)	(-0.1%)	(-2.0%)	(-2.0%)	(-1.3%)	(-1.1%)	(-6.3%)
10	142	186,261	186,061	186,173	182,354	182,420	183,828	183,696	172,719
			(-0.1%)	(-0.0%)	(-2.1%)	(-2.1%)	(-1.3%)	(-1.4%)	(-7.3%)
11	134	172,300	172,232	172,290	168,347	168,430	170,054	168,669	156,131
			(-0.0%)	(-0.0%)	(-2.3%)	(-2.2%)	(-1.3%)	(-2.1%)	(-9.4%)
12	155	181,025	180,928	180,363	177,081	177,160	178,746	177,593	165,323
			(-0.1%)	(-0.4%)	(-2.2%)	(-2.1%)	(-1.3%)	(-1.9%)	(-8.7%)
13	164	170,543	170,377	170,540	166,549	166,648	168,448	165,977	152,311
			(-0.1%)	(-0.0%)	(-2.3%)	(-2.3%)	(-1.2%)	(-2.7%)	(-10.%)
14	147	177,945	177,897	177,940	173,983	174,069	175,732	174,113	161,318
			(-0.0%)	(-0.0%)	(-2.2%)	(-2.2%)	(-1.2%)	(-2.2%)	(-9.3%)
15	140	172,311	172,304	172,304	168,308	168,410	170,245	167,555	153,684
			(-0.0%)	(-0.0%)	(-2.3%)	(-2.3%)	(-1.2%)	(-2.8%)	(-10.%)
16	131	149,788	149,774	149,743	145,594	145,781	148,326	140,746	123,583
			(-0.0%)	(-0.0%)	(-2.8%)	(-2.7%)	(-1.0%)	(-6.0%)	(-17.%)

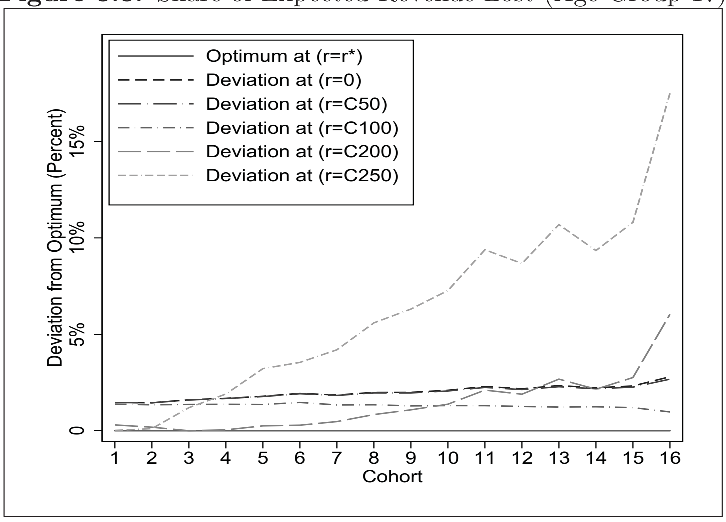
Notes: Deviations from the maximum expected revenue are stated in parentheses. Since largely identical to the respective median reserve price (see Table 3.5), the results for cluster point C_{150} are omitted from the presentation.

magnitude. Returning to Table 3.5, note that all of these values are located considerably below r_j^* for $j \in \{1, ..., 16\}$. As a consequence, at these levels the reserve price is likely to be non-binding in the sense that the odds for at least two of the assumed $N = 7$ potential bidders realizing a higher valuation are reasonably large. Intuitively, these sellers forego the chance to gain additional rents from setting a reserve price that lies above the expected second-highest bidder's valuation. The auction outcome is thus effectively determined by the degree of competition among the potential bidders.

A higher number of potential bidders implies a larger number of i.i.d. draws from the valuation primitive. Since this ceteris paribus increases the probability for the realization of high valuations, competition is more intense for higher N. Consistently, in the estimation for $N = 13$ and $N = 18$ the shares of expected revenue lost relative to the optimum are of even smaller magnitude than for $N = 7$. Conversely, for $N = 2$ potential bidders also the expected revenues at zero, C_{50}, and C_{100} dramatically deviate from the optimum, since the compensating impact of bidder competition is substantially reduced. Since HT's transfer market is highly competitive, our intuition is that the assumption of at least $N = 7$ potential bidders is indeed justified and that the finding of a relatively small loss in expected revenue due to a suboptimally low reserve price adequately describes the actual situation.

As soon as the minimum bid exceeds the optimal reserve price, however, the shares of foregone expected revenue increase substantially. Intuitively, an exaggerated minimum bid induces an inefficiently high probability that the auction fails, which causes the expected revenues for the seller to decline. For instance, the optimal reserve price in cohort $j = 9$ is given by $r_9^* = 161,678$. Yet, 15% of the sellers charged an reserve price equal to or above C_{250}, thereby accruing a loss of more than 6.3% relative to the optimum. Moreover, the more the chosen minimum bid exceeds r_j^*, the larger is the share of expected revenue lost. In cohort $j = 16$, in 10% of the 131 auctions the seller demanded a starting price of C_{250} or higher, thereby reducing his expected revenue by more than 17% relative to the optimum. Figure 3.8 illustrates the respective deviations for all 16 cohorts in age-group seventeen.[42] Clearly, the shares of expected revenue lost at minimum bids at levels of C_{200} and C_{250} are amplified for older players in the age-group, since the respective optimal reserve prices decline as age increases and thus are exceedingly outvalued. Consistent with the above argument that bidder competition triggers a compensating effect, for reserve prices below r_j^*, i.e. $r_j \in \{0, C_{50}, C_{100}\}$, the measured deviations remain relatively constant.

Figure 3.8: Share of Expected Revenue Lost (Age-Group 17)



[42] For an analogous illustration for age-groups eighteen and nineteen see Figure 3.11 in Appendix 3.6.

Result 6 (Loss of expected revenue): *The impact of a suboptimal reserve price depends on the direction of the deviation from its optimal level. Downward deviations, i.e. setting a non-binding reserve price, are partially countervailed by the fact that competition among the potential bidders is sufficiently intense. In contrast, upward deviations can cause substantial reductions in expected revenue.*

Summarizing, though suboptimal from a theoretical perspective, the large number of zero reserve prices and cluster points we observe in our data not necessarily trigger severe deviations from the optimum. As long as the reserve price is set below its optimal level, the competitive environment of the transfer market suffices to guarantee an almost optimal expected revenue for the seller, though he will not succeed to additionally extract some of the winning bidder's surplus. In contrast, a large fraction of sellers loses a substantial share of expected revenue by setting a reserve price too high, thereby increasing the probability that the player remains unsold.

3.5 Conclusion

We examine empirically how managers playing the online game Hattrick set reserve prices in auctions for virtual players. Using detailed field data on 6,258 auctions from *HT*'s transfer market, we find that chosen reserve prices exhibit both, very sophisticated and suboptimal behavior by the sellers. Reserve prices pick up the birthday effect of the demand side documented in Chapter 2 and are adjusted remarkably nuanced to the resulting sales price pattern. All our results are robust if we additionally control for the experience of sellers, the auction-end-day, and time-of-day effects.

Intriguingly, even though *HT*'s transfer market is highly competitive, we find evidence for a sunk cost fallacy and a resulting entitlement effect in form of a large positive premium on the reserve price when a player has been acquired previously.

While many sellers act strategically and try to reap some of the buyers' surplus, some fail in this endeavor as they set reserve prices suboptimally. We have established that reserve prices are too clustered (around € 50,000 steps) to be consistent with fully rational behavior and we document what share of expected revenues is foregone by this. We thus conclude that many *HT* managers simplify their decision making by adopting heuristic pricing rules that are suboptimal from a fully rational point of view.

If, as in our data, the entitlement and clustering effects are persistent and quantitatively relevant, the option of choosing a reserve price might be an impediment to market efficiency as sellers set too high reserve prices resulting in too little trade. In such situations a social planner might want to avoid using a reserve price in the design of an auction format to prevent this potential distortion.

On the upside, our findings suggest that simple microeconomic theory gives us a lot of mileage in explaining market behavior in complex environments. We document that (the majority of) sellers very finely adjust their behavior to demand patterns and try to strategically exploit potential arbitrage possibilities. Moreover, we are able to show that the adoption of heuristic pricing rules by the sellers does not affect the expected revenue from an auction dramatically, as long as the chosen reserve price lies below the optimal level and competition among the bidders is sufficiently intense.

Clearly, more research on the behavior of sellers under different auction formats is needed to improve our understanding of the determinants and effects of reserve prices and to evaluate the efficiency of different market designs. Much of the existing evidence derives from field data or field experiments. Hence, it might be worthwhile to conduct controlled laboratory experiments with focus on the supply side of auction markets. They allow for a systematic variation of the design features that might possibly drive the behavioral patterns observed in the field, while at the same time resolving many factors of uncertainty like the lack of knowledge of the underlying valuation primitive. Alternatively, as a way of bridging the gap between the lab and the field, virtual economies like *HT* may also serve as platforms to conduct controlled economic and social experiments. Although considerably more complex than most laboratory experiments, and despite the missing monetary incentives, the findings presented in this and other studies (see e.g. Trautmann and Traxler, 2009, Castronova, 2008, and Nicklisch and Salz, 2008) suggest that they provide a market framework which adheres to standard economic constraints while still providing a considerable degree of control.

3.6 Appendix

Table 3.7: Determinants of Price Unconditional on Sale (Tobit)

Dep.Var.: *price*	Coeff.	Std. Dev
days17	-1609.81***	64.72
days18	-138.30**	56.70
days19	-10.70	49.68
age18	-173269.62***	5834.23
age19	-196106.93***	5644.17
tsi	88.22***	4.75
acquired	-13578.76***	3203.95
Intercept	79122.68***	16476.41
skills	yes	
character	yes	
daytime	yes	
weekday	yes	
LR Chi(28)	2587.09	
N (1515 left-cens. at 0)	6258	

Notes: The TOBIT procedure also includes failed auctions with a sales price of zero (left-censored). Standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level.

Table 3.8: Expected Revenues for Age-Group 18 (Cohorts 17-32)

		I	II	III	IV	V	VI	VII
Coh.	Obs.	$\tilde{\Pi}(r_j^*)$	$\tilde{\Pi}(r_j^{mean})$	$\tilde{\Pi}(r_j^{med})$	$\tilde{\Pi}(0)$	$\tilde{\Pi}(C_{50})$	$\tilde{\Pi}(C_{100})$	$\tilde{\Pi}(C_{150})$
17	102	113,912	111,070	110,433	109,007	109,656	113,710	109,149
			(-2.5%)	(-3.1%)	(-4.3%)	(-3.7%)	(-0.2%)	(-4.2%)
18	125	94,990	92,663	91,217	89,365	90,691	94,950	85,867
			(-2.5%)	(-4.0%)	(-5.9%)	(-4.5%)	(-0.0%)	(-9.6%)
19	111	94,375	91,187	91,173	88,780	90,075	94,344	85,405
			(-3.4%)	(-3.4%)	(-5.9%)	(-4.6%)	(-0.0%)	(-9.5%)
20	116	88,289	84,047	84,025	82,346	84,025	88,089	77,669
			(-4.8%)	(-4.8%)	(-6.7%)	(-4.8%)	(-0.2%)	(-12.%)
21	119	94,140	89,874	89,841	88,517	89,841	94,101	85,025
			(-4.5%)	(-4.6%)	(-6.0%)	(-4.6%)	(-0.0%)	(-9.7%)
22	123	87,899	85,079	83,643	81,905	83,643	87,664	77,068
			(-3.2%)	(-4.8%)	(-6.8%)	(-4.8%)	(-0.3%)	(-12.%)
23	125	86,968	82,985	82,722	80,918	82,722	86,692	75,909
			(-4.6%)	(-4.9%)	(-7.0%)	(-4.9%)	(-0.3%)	(-12.%)
24	141	99,835	96,692	95,533	94,534	95,533	99,830	92,529
			(-3.1%)	(-4.3%)	(-5.3%)	(-4.3%)	(-0.0%)	(-7.3%)
25	115	85,880	81,979	81,645	79,771	81,645	85,559	74,592
			(-4.5%)	(-4.9%)	(-7.1%)	(-4.9%)	(-0.4%)	(-13.%)
26	117	89,884	86,234	85,611	84,001	85,611	89,722	79,524
			(-4.1%)	(-4.8%)	(-6.5%)	(-4.8%)	(-0.2%)	(-11.%)
27	104	81,142	77,587	76,950	74,838	76,950	80,653	69,175
			(-4.4%)	(-5.2%)	(-7.8%)	(-5.2%)	(-0.6%)	(-14.%)
28	105	82,464	78,741	78,280	76,126	78,280	81,944	70,390
			(-4.5%)	(-5.1%)	(-7.7%)	(-5.1%)	(-0.6%)	(-14.%)
29	124	86,198	83,066	82,254	80,191	81,944	85,954	75,314
			(-3.6%)	(-4.6%)	(-7.0%)	(-4.9%)	(-0.3%)	(-12.%)
30	136	81,068	76,603	76,890	74,710	76,890	80,529	68,930
			(-5.5%)	(-5.2%)	(-7.8%)	(-5.2%)	(-0.7%)	(-15.%)
31	139	82,104	78,925	77,932	75,719	77,932	81,539	69,886
			(-3.9%)	(-5.1%)	(-7.8%)	(-5.1%)	(-0.7%)	(-14.%)
32	131	74,430	70,400	70,180	67,772	70,336	73,587	61,488
			(-5.4%)	(-5.7%)	(-8.9%)	(-5.5%)	(-1.1%)	(-17.%)

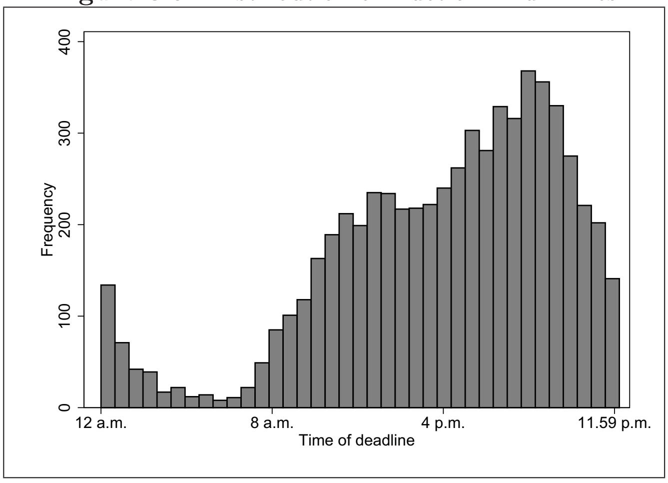
Notes: Deviations from the maximum expected revenue are stated in parentheses.

Table 3.9: Expected Revenues for Age-Group 19 (Cohorts 33-48)

		I	II	III	IV	V	VI
Coh.	Obs.	$\tilde{\Pi}(r_j^*)$	$\tilde{\Pi}(r_j^{mean})$	$\tilde{\Pi}(r_j^{med})$	$\tilde{\Pi}(0)$	$\tilde{\Pi}(C_{50})$	$\tilde{\Pi}(C_{100})$
33	160	73,387	68,706	67,818	66,271	69,454	72,032
			(-6.4%)	(-7.6%)	(-9.7%)	(-5.4%)	(-1.8%)
34	165	72,115	66,423	66,506	64,729	68,289	70,451
			(-7.9%)	(-7.8%)	(-10.%)	(-5.3%)	(-2.3%)
35	151	64,982	59,187	58,430	56,885	61,477	62,535
			(-8.9%)	(-10.%)	(-12.%)	(-5.4%)	(-3.8%)
36	144	69,869	64,070	64,167	62,360	66,096	68,065
			(-8.3%)	(-8.2%)	(-10.%)	(-5.4%)	(-2.6%)
37	127	64,422	58,422	57,873	56,337	60,911	61,987
			(-9.3%)	(-10.%)	(-12.%)	(-5.4%)	(-3.8%)
38	118	68,667	62,415	62,218	60,893	65,011	66,566
			(-9.1%)	(-9.4%)	(-11.%)	(-5.3%)	(-3.1%)
39	138	71,575	65,288	65,836	64,306	67,702	70,045
			(-8.8%)	(-8.0%)	(-10.%)	(-5.4%)	(-2.1%)
40	139	68,075	62,068	62,023	60,377	64,386	66,058
			(-8.8%)	(-8.9%)	(-11.%)	(-5.4%)	(-3.0%)
41	175	73,928	68,342	68,208	66,746	70,021	72,498
			(-7.6%)	(-7.7%)	(-9.7%)	(-5.3%)	(-1.9%)
42	152	66,852	61,839	60,371	58,983	63,240	64,647
			(-7.5%)	(-9.7%)	(-11.%)	(-5.4%)	(-3.3%)
43	138	70,707	64,875	65,077	63,113	66,971	68,806
			(-8.2%)	(-8.0%)	(-10.%)	(-5.3%)	(-2.7%)
44	165	64,965	59,139	58,816	56,673	61,554	62,321
			(-9.0%)	(-9.5%)	(-12.%)	(-5.3%)	(-4.1%)
45	186	68,156	62,728	61,705	60,374	64,504	66,046
			(-8.0%)	(-9.5%)	(-11.%)	(-5.4%)	(-3.1%)
46	169	64,349	57,849	57,868	56,480	60,736	62,144
			(-10.%)	(-10.%)	(-12.%)	(-5.6%)	(-3.4%)
47	137	63,836	57,752	58,187	55,721	60,340	61,370
			(-9.5%)	(-8.9%)	(-12.%)	(-5.5%)	(-3.9%)
48	171	68,474	62,479	62,021	60,688	64,823	66,360
			(-8.8%)	(-9.4%)	(-11.%)	(-5.3%)	(-3.1%)

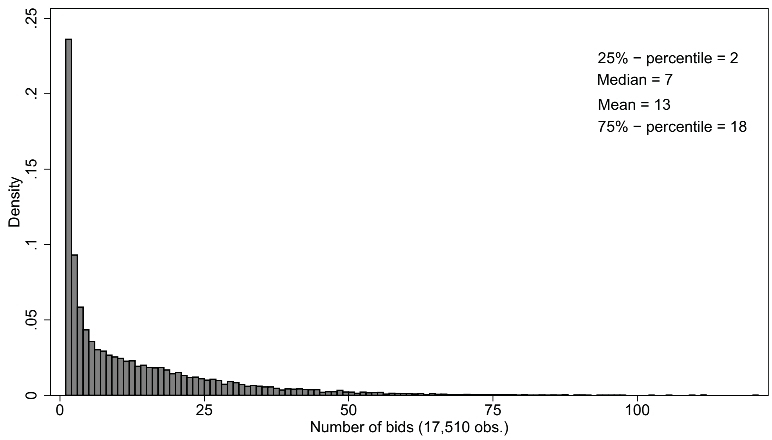
Notes: Deviations from the maximum expected revenue are stated in parentheses.

Figure 3.9: Distribution of Auction End-Times



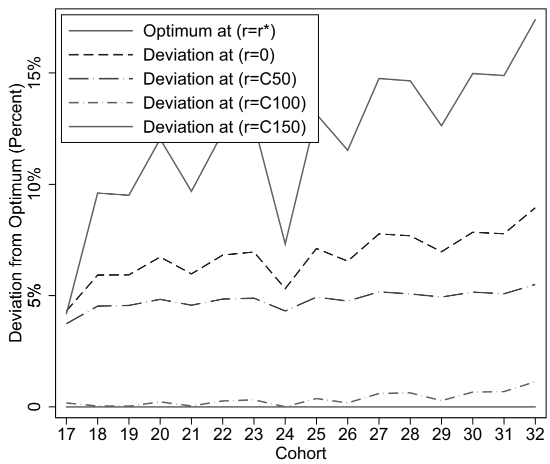
Notes: During a typical day, the time of the auction deadlines is approximately similarly distributed as the number of simultaneously logged-on users. During the early morning hours CET we observe the lowest traffic with roughly 7,000 online users at its minimum, while during the peak-periods between 5:30 p.m. and 10 p.m. levels up to 75,000 simultaneously online users are reached.

Figure 3.10: Distribution of the Number of Bids

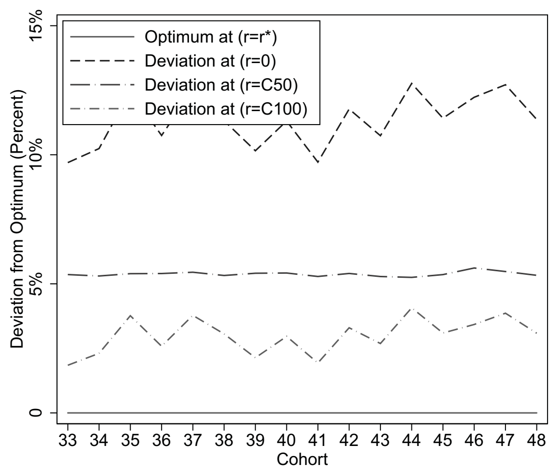


Notes: The graph depicts the distribution of the number of bids observed in the sample of 17,510 Hattrick-auctions that was used in Chapter 2. This information is employed as a proxy for the number of bidders in the estimation of the valuation primitive.

Figure 3.11: Share of Expected Revenue Lost for Age-Groups 18 and 19



(a) Age-Group 18



(b) Age-Group 19

Chapter 4

The Evaluation of Complex Goods - Evidence From Online Car Sales

4.1 Introduction

Modern consumer goods are often complex in nature in the sense that they consist of a multitude of characteristics that collectively account for their market value. With increasing frequency people are challenged to evaluate goods like cars, mobile phones, or personal computers on the basis of their various constituting attributes. For instance, the value of a laptop depends on the processor speed, the capacity of the hard-drive, the graphics engine, and a lot of other features, which all have to be individually accounted for as they determine its overall quality. While economic theory suggests that a rational agent should incorporate all relevant pieces of information into his considerations and exclude any that are non-informative, in the past such an evaluation task was often made difficult by the lack of accessible sources to gather the necessary information. With the rise of the internet, however, today a plethora of valuable information is often just a few clicks away. A large number of online platforms specialized on specific complex goods, e.g. used cars, makes it possible to easily cross-compare similar offers for close substitute objects. Therefore, we are especially interested in the behavior of individuals when they have to perform this task: Do people efficiently identify and incorporate all of the pertinent elements, or are some valuable pieces of information not reflected in their individual evaluation of the good?

Based on detailed field data on used car offers from *mobile.de*, one of Europe's largest online vehicle marketplaces, we address this question empirically. We conduct a hedonic regression analysis to identify which of the cars' features significantly influence the stated sales prices and if their impact stands in accordance to the relation we would rationally expect. For instance, since the odometer reading conveys an imperfect signal for the abrasion of engine and chassis components, the mileage should be negatively correlated with a car's value, while the existence of extras, e.g. a sun-roof or an automatic gearbox should have a positive impact on the sales price. By this approach, we are able to examine the extent to which individual evaluations of used cars efficiently reflect all relevant and openly accessible pieces of information.

Despite the fact that many effects are consistent with our predictions, our data shows that people are systematically inattentive to a crucial piece of valuable information, even if it is readily provided. In particular, though both the precise month and year of first registration (FR) are publicly and prominently stated for each car, we find highly significant differences of sizeable magnitude for the impact of a marginal month on the sales price depending on whether a car was registered in the previous or in the same year. All else equal, the price differential between two cars, where one was first registered in January and the other in December of the previous year, is dramatically larger than that between two cars first registered in any two subsequent months of the same year, respectively. Stated differently, we find an amplified adjustment in the prices for otherwise identical cars to be located *across* different registration years, or "vintages", where the impact of a marginal month of age is *up to four times larger* relative to that *within* the same vintage. As a consequence, the observed price pattern evolves non-continuously and exhibits distinct discontinuities where the registration year changes. Our intuition is that individuals over-react to the figure displayed in the registration year, i.e. the "vintage", while they are inattentive to the finer information as conveyed through the month of first registration. In other words, a car systematically loses in value simply for the fact that it displays a different digit in the count of its registration year compared to another that is factually just one month younger. This result remains robust across all four different makes and models we analyze and also across various specifications and additional controls.

Our findings strongly complement the results of Englmaier and Schmöller (2009b), as presented in Chapter 2 of this dissertation. Similarly to the approach taken in this paper, there we examine to what extent individual bidding behavior reflects valuable information that is readily provided by analyzing auction prices paid for virtual football players in the online game Hattrick (*HT*). Importantly, these players closely resemble complex goods as they are completely described by a multi-dimensional vector of attributes that collectively determine their value. While the underlying game algorithm strongly suggests a continuous relation between the age of a player and his value, akin to the effect of the registration year on the price of a used car, we find evidence for a large systematic drop in sales prices just on a player's birthday, a phenomenon which we therefore label the "birthday effect". Intuitively, this indicates that buyers in *HT* give too much weight to the age of a player measured in years as opposed to his age measured in days, though the latter is also plainly visible to all buyers free of cost.

Although the market environment in the game provides considerable incentives for the users when engaging on the transfer market, all transactions in *HT* are carried out in terms of virtual money. In this study, we therefore test for the external validity of the "birthday effect" documented in the virtual *HT*-economy by analyzing to what extent information is efficiently used in a market that involves large real monetary stakes on part of the trading parties, i.e. the purchase and sale of used cars.

Other than *HT*, our present sample of used cars does not originate from an auction market, and the economic environment of *mobile.de* is considerably less controlled than the highly structured transfer market of *HT*. In addition, rather than on actual sales prices, our analysis is based on the asking prices stated by the individual sellers, which may be subject to negotiation once an interested buyer has been found. However, we have strong reasons to believe that the posted price is a sensible proxy for the final price in this market. First, *mobile.de* offers the seller an option to declare the stated price either as "fixed" or as "negotiable", and a substantial fraction of the sellers opts for the former rather than the latter. Second, with several thousand offers for each model series the market for used cars is highly competitive. Moreover, the cars within each of our subsamples can be regarded as close substitutes. Under the presumption, that the stated sales price reflects the willingness to accept of the respective seller, according to Hanemann (1991) and Shogren et al. (1994)

in such an environment an endowment effect, i.e. a divergence of willingness to pay and willingness to accept, is unlikely to persist.[1] Hence, it stands to reason that the stated prices are considerably close to the final prices. Finally, since advertising a car is costly, it seems plausible that the sellers exert considerable effort to elicit a reasonable price, at which prospective buyers are indeed willing to buy.[2] For simplicity, in the following we use the term "price" to refer to the stated prices in our data.

Under these premises, the basic situation on *mobile.de* is comparable to that of the buyers in *HT*: People are presented with many details of a complex good and have to form their valuation for it. Moreover, the virtual players in the game closely resemble the used cars in our present data in several respects. First, both are traded in large numbers on specialized internet market platforms. Second, they constitute complex goods that can be decomposed into their constituent characteristics, for each of which we can obtain estimates of the contributory value using a hedonic regression approach. Third, in both cases we are able to create subsamples of close substitutes by focusing on a particular model series or player type, respectively. Fourth, similar to the registration date of used cars, the age of a virtual player is displayed through two dimensions, the year and days of age. Finally, and most importantly, analogously to the virtual players, ceteris paribus a car can be rationally expected to depreciate continuously in value as its age increases.

But there is also an important difference: Buying a car constitutes a major purchase for most households and thus should be subject to profound pricing considerations. Yet, similar to the buyers in *HT*, we find that people systematically underrate the information on the precise age, i.e. the month of first registration, though explicitly provided. Thus, given the large monetary stakes involved, the fact that we observe congeneric discontinuities in the price pattern for used cars not only substantiates the external validity of the "birthday effect", but also shows that such an evaluation bias can have real economic consequences.

Inattentiveness and limited attention have also been documented for other purchase decisions in other markets. For instance, Lee and Malmendier (2007) analyze individual bidding behavior in auctions on eBay and find that people tend to anchor on an irrelevant outside

[1] Moreover, the services of *mobile.de* are widely used by professional car dealers who purchase cars for resale rather than use, where according to Kahneman et al. (1991) the endowment effect does not apply. As it turns out, the majority of offers in our sample is indeed made by commercial rather than private sellers.

[2] In line with this argument, in Englmaier and Schmöller (2009a) (see Chapter 3) we document that the sellers' reserve prices in *HT*-auctions are similarly determined as the sales prices, i.e. from an evaluation of the individual attributes. Our intuition is that the same also applies to a non-auction context.

retail price for a board game, if the seller chose to state that price in the description of the product details. At the same time, many of the winning bids exceed a more relevant outside option, the so called "buy-it-now" price, which is an ex-ante fixed strike price set by the seller as an alternative to the auction process. Analyzing stock market data, Gilbert et al. (2008) provide evidence that investors with limited attention have an incentive to focus on summary statistics rather than individual pieces of information. They analyze the market response to the U.S. Leading Economic Index (LEI), a macroeconomic release that is purely a summary statistic, and show that the LEI announcement has an impact on aggregate stock returns, return volatility, and trading volume. We add to these findings by demonstrating that inattentiveness effects pertain for complex goods and large stake purchase decisions, even though the concerned piece of information is provided at arm's length within the relevant market environment.

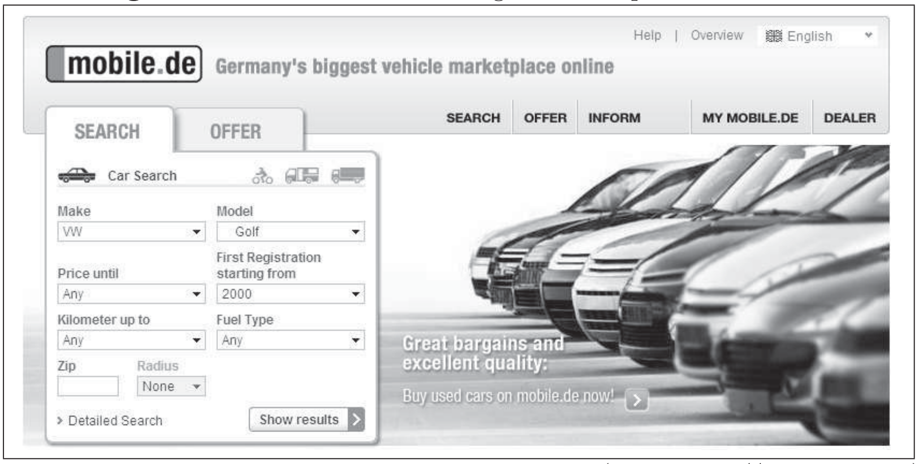
The remainder of the paper is structured as follows. Section 4.2 describes the structure of the data and the relevant details of the sample selection. Section 4.3 presents the details of our empirical estimation and the results from the hedonic regression analysis. In addition, we provide a series of robustness checks of our findings and briefly discuss the results from an alternative estimation aproach. Section 4.4 concludes and an Appendix collects additional Tables and Figures.

4.2 Data Description

4.2.1 Institutional Background

For the purpose of this study, we collected detailed information on more than 80,000 cars offered during July and August 2009 on the online vehicle market platform *mobile.de*. Founded in 1996, *mobile.de* takes the role of an intermediator between supply and demand within a two-sided market. The company itself is not involved at any stage in the purchase or sale of a vehicle and a successful sale does not invoke any final value fees to *mobile.de*. It provides both a platform for sellers to place advertisements for new and used cars at a small cost and a free comprehensive search tool for prospective buyers to screen among the mass of on average about 1.3 million offers. According to the company's own statement, prospective buyers "*can limit search results by setting individual preferences and like this obtain customized offers with just a few clicks*", providing them "[...] *with an overview of the market and information about prices"*.[3] Hence, the same is true for a seller who wants to evaluate his car before placing a sales advertisement.

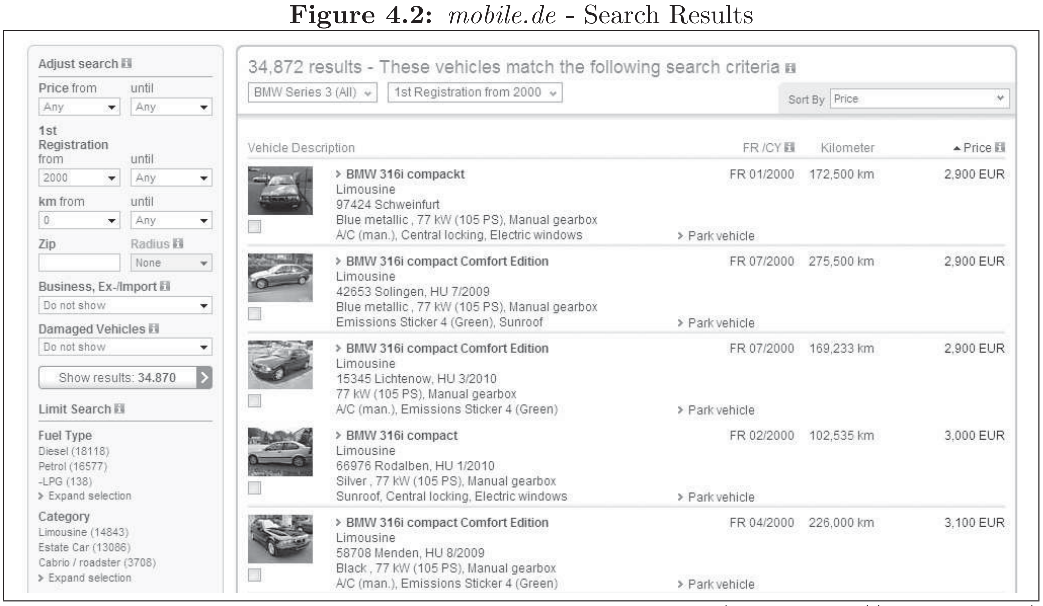
Figure 4.1: *mobile.de* - Start Page with Simple Search Form



(Source: http://www.mobile.de)

Figure 4.1 shows the interface a user is presented with upon entering *mobile.de*'s website. It displays a simple search form, which among other things allows to filter for makes, models, and a number of other basic details. A detailed search form, which can be directly reached by clicking the link to the lower left, provides a large additional set of filter options. Note however, that the drop down selector for the date of first registration only allows to filter for the vintage, i.e. the FR-year. Neither in the simple nor the expanded form it is possible to adjust the search inquiry for the precise month of the first registration (FR).

Figure 4.2: *mobile.de* - Search Results



(Source: http://www.mobile.de)

[3]Source: http://cms.mobile.de/en/company/portrait_mobile.html

The search returns a list of all vehicles matching the chosen filters. Per default they are sorted by price, where an abstract of their main features is displayed as shown in Figure 4.2. This preview explicitly states the precise date of first registration (e.g. "FR 01/2000") and additionally provides valuable information on the price, mileage, color, and power of the car, to name only a few. It is also possible to remember a specific car for later access ("Park vehicle"), which allows the user to directly compare the latter to other remembered cars.

A typical profile page of an offered car, which is accessed from the search results list by clicking on the model name at the top of the respective entry, is depicted in Figure 4.3. So far, the described environment is virtually identical to the one analyzed in Englmaier and Schmöller (2009b). Unlike in *HT*, however, the process of buying is not carried out on *mobile.de* directly, but rather a prospective buyer is merely provided with the contact details of the respective seller. Moreover, the product descriptions are composed by the individual sellers and not fully standardized as in *HT*'s transfer market.

Figure 4.3: *mobile.de* - Vehicle Profile

(Source: http://www.mobile.de)

For each car, a seller has to specify a preselected set of features and attributes, where most of the respective values are chosen from a drop down menu during the preparation of the advertisement. Conveniently, this data is thus standardized and ensures a sufficient degree of comparability across individual observations. Naturally, we therefore focus on these standard attributes in our data, which in addition to the stated price and the date of FR include various extras and also some information on the sellers (see Table 4.2).[4]

Having described the environment our data stems from, we next turn to a detailed discussion of our sample selection criteria.

4.2.2 Sample selection

Our data includes details on the most widespread car models from four leading German makes, all ranked among the top seven of Germany's vehicle population according to the Kraftfahrt-Bundesamt (KBA).[5] More specifically, we collected information on 29,097 Volkswagen (VW) Golf (KBA-rank 1), 14,693 Opel Astra (KBA-rank 2), 25,582 BMW 3 (KBA-rank 4), and 17,901 Audi A4 (KBA-rank 7), all advertised as accident-free and with their FR-dates between 01/2000 and 12/2008.[6] We focus on this subsample for two main reasons. First, a high stock is a good indicator for a considerable volume of used car offers for a specific model, which ensures a sufficiently large number of observations. Second, we consider models from different makes to achieve a broad diversification within our identification strategy.[7]

Since the introduction of a new series within a particular car model affects the sales prices substantially, we can only retrieve meaningful estimates of the influential attributes if we accurately control for potential model revisions. Clearly, this requires detailed knowledge of the exact dates of the respective market launches. Conveniently, for the four different models considered in our sample, this information is readily available. In particular, we identify the respective estimation windows for each model according to the information provided through the manufacturers' websites, the Schwacke-List (http://schwacke.de), and

[4] Any additional information provided by a seller takes a free text form, which would require us to manually convert these into a standardized format to be able to employ them for the analysis. However, the set of features included within our sample is already quite comprehensive and suffices to explain much of the variation observed in the stated prices.

[5] Source: http://www.kba.de.

[6] KBA-ranks not reported were taken by other models of VW (Passat, Polo) and Opel (Corsa).

[7] We do not consider cars that were first registered before January 2000, since their values are very low. Moreover, we thereby avoid a potential "left-digit" effect with respect to the registration year: It has been documented that some individuals tend to process numerical information in a way that the first digits are treated as more valuable information, i.e. are perceived to contain more significant information than later digits (see e.g. Brenner and Brenner, 1982; Bhattacharya et al., 2008).

the Deutsche Automobil Treuhand (http://www.dat.de).[8] Since all models in consideration experienced at least one upgrade or change of series between 2000 and 2008, we define the estimation periods accordingly and are thus able to conduct our analysis for eight different subsamples, as shown in Table 4.1.[9]

Table 4.1: Models Series and Estimation Periods

Make & Model	Name of Series		Production period	Estimation period
VW Golf	IV		10/1997 ~ 09/2003	01/2000 - 09/2003
	V		10/2003 ~ 07/2008	01/2004 - 06/2008
BMW 3	E46		04/1998 - 11/2004	01/2000 - 11/2004
	E90	(limousine)	12/2004 - 09/2008*	09/2005 - 08/2008
		(estate)	06/2005 - 09/2008*	
Audi A4	B6	(limousine)	10/2000 - 11/2004	10/2001 - 10/2004
		(estate)	09/2001 - 11/2004	
	B7	(limousine)	11/2004 - 11/2007	04/2005 - 09/2007
		(estate)	11/2004 - 03/2008	
Opel Astra	G		02/1998 - 01/2004	01/2000 - 12/2003
	H		02/2004 - 10/2007*	05/2004 - 10/2007

Notes: Entries with an asterisk indicate an upgrade of the current production series. If there were different introduction dates within a model series, we use the later date to determine the estimation period.

After each change of series, we include a short transition period before the respective estimation window, to ensure that the cars within each subsample belong to the same series of a model, i.e. can plausibly be perceived as close substitutes. Naturally, due to different variants offered within a model series, e.g. limousine, estate car, or compact car, the latter are no perfect substitutes. To account for such within-series variation, we distinguish between five- and three door versions, add a large set of main attributes as controls, and exclude convertibles from the sample. In this way, we capture a substantial share of the variation in the price within a series and are thus able to obtain precise estimates of the influential factors.

Since a complete description of all eight subsamples would go beyond the scope of the paper, throughout the following we representatively focus on the samples of VW Golf series IV and V, and provide the corresponding details for the other three models in Appendix 4.5. In all cases, the analysis yields very similar and qualitatively robust results.

[8]The latter are commercial service providers who offer benchmark evaluations for all kind of cars at a small cost. In fact, they allow to account for the precise date of first registration in an individual evaluation of a car, which makes the discontinuities we are able to document in our data even more puzzling.

[9]Depending on their extent, these upgrades, or “face-lifts”, can invoke similar price effects as a change of series. If available, in the estimation we therefore treat the information on a face-lift similar to the introduction of a new production line.

4.2.3 Data description

In general, the value of an individual car from a specific model series depends on numerous factors. Among others, this includes its age, its odometer reading, the power and fuel-type of its engine, and the different extras it is equipped with, e.g. an automatic gearbox, a sun-roof, a seat-heating, or a cruise control. Along with the stated prices and the month and year of first registration, we therefore collected a large number of features for each of the cars to control directly for quality differences. To measure their impact on the price of the car, we assign a dummy variable to each of the observed extras in our analysis. For instance, if a offered car has a sun-roof, the dummy *sun_roof* takes value 1 and 0 otherwise.[10] Similarly, we also add a dummy for both the door-count and the fuel type.[11] Table 4.2 provides an overview of the collected details and shows the corresponding summary statistics for the samples of VW Golf series IV and V, respectively.[12]

Naturally, the restriction to the estimation windows as described above implies that 6,807 of the overall observations in the subgroup of VW Golf are not considered for the analysis, reducing the sample size to 22,290. In addition, we drop all entries with missing-values for one or more of the considered variables and correct for outliers with respect to *mileage* and *price*.[13] This leaves us with a final sample of 6,034 and 15,247 observations in series IV and V, respectively, or 95% of the initial data points within the relevant estimation periods.

The information on the month and year of the first registration is stored in the variables *fr_month* $\in [1, 12]$ and *fr_year* $\in [2000, 2008]$, respectively. For our empirical analysis, we combine the latter to construct the measure *totalage* $\in [1, 108]$, which displays the precise age of a car in units of months:

$$\textit{totalage} \equiv 12 \cdot (2008 - \textit{fr\_year}) + (13 - \textit{fr\_month}),$$

where the normalization is such that a car's age is measured relative to the most recent FR-date included within our dataset, i.e. 12/2008, which corresponds to the minimum age of 1 month. Analogously, for cars with an FR-date of 01/2000, i.e. the oldest cars in our sample, *totalage* takes its maximum at 108 months.

[10] In the following, we use italics to denote the variable name in our data corresponding to an attribute.

[11] For air conditioning, airbags, and electric window lifters we find almost no variation in the data. Since by now these features are included in the basic configuration of most cars, we omit them from the analysis.

[12] The corresponding tables for BMW 3, Audi A4 and Opel Astra are provided in Appendix 4.5.

[13] Outliers are classified as values above the respective 99th-percentile in each series. None of the results depends on their omission.

Table 4.2: Summary Statistics - VW Golf

	Panel A. Series IV (Est. Period: 01/2000 - 09/2003)				Panel B. Series V (Est. Period: 01/2004 - 06/2008)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	6,034	6,711	2,499	17,990	15,247	13,605	4,900	27,979
mileage [km]	6,034	115,978	2,000	225,500	15,247	55,343	1,000	221321
power [kW]	6,034	73	50	213	15,247	81	50	243
fr_year	6,034	2001	2000	2003	15,247	2006	2004	2008
fr_month	6,034	6	1	12	15,247	6	1	12
totalage [months][a]	6,034	87	64	108	15,247	30	7	60
Dummies[b]	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	3,771	62.5	62.5	0	5,998	39.3	39.3
(0 = petrol, 1 = diesel)	1	2,263	37.5	100.0	1	9,249	60.7	100.0
five-door	0	1,656	27.4	27.4	0	2,921	19.2	19.2
(0 = three, 1 = five)	1	4,378	72.6	100.0	1	12,326	80.8	100.0
auto gearbox	0	5,509	91.3	91.3	0	13,364	87.7	87.7
(0 = manu., 1 = auto)	1	525	8.7	100.0	1	1,883	12.4	100.0
cruise control	0	5,222	86.5	86.5	0	7,895	51.8	51.8
(0 = no, 1 = yes)	1	812	13.5	100.0	1	7,352	48.2	100.0
seat heating	0	4,554	75.5	75.5	0	7,712	50.6	50.6
(0 = no, 1 = yes)	1	1,480	24.5	100.0	1	7,535	49.4	100.0
all-wheel drive	0	5,879	97.4	97.4	0	7,712	50.6	50.6
(0 = no, 1 = yes)	1	155	2.6	100.0	1	7,535	49.4	100.0
sun roof	0	5,106	84.6	84.6	0	13,634	89.4	89.4
(0 = no, 1 = yes)	1	928	15.4	100.0	1	1,613	10.6	100.0
leathertrim	0	5,817	96.4	96.4	0	14,560	95.5	95.5
(0 = no, 1 = yes)	1	217	3.6	100.0	1	687	4.5	100.0
metallic paint	0	1,589	26.3	26.3	0	3,258	21.4	21.4
(0 = no, 1 = yes)	1	4,445	73.7	100.0	1	11,989	78.6	100.0
private seller	0	4,419	73.2	73.2	0	13,821	90.7	90.7
(0 = no, 1 = yes)	1	1,615	26.8	100.0	1	1,426	9.4	100.0

a *totalage* $\equiv$ 12·(2008-fr_year)+(13-fr_month) displays the age of a car in units of months and is normalized such that the minimum 1 $\widehat{=}$ 12/2008 and the maximum 108 $\widehat{=}$ 01/2000. For instance, the estimation period for VW Golf IV, i.e. 01/2000 to 09/2003, equals *totalage* $\in [64, 108]$. **b** Color dummies are not displayed to save on space.

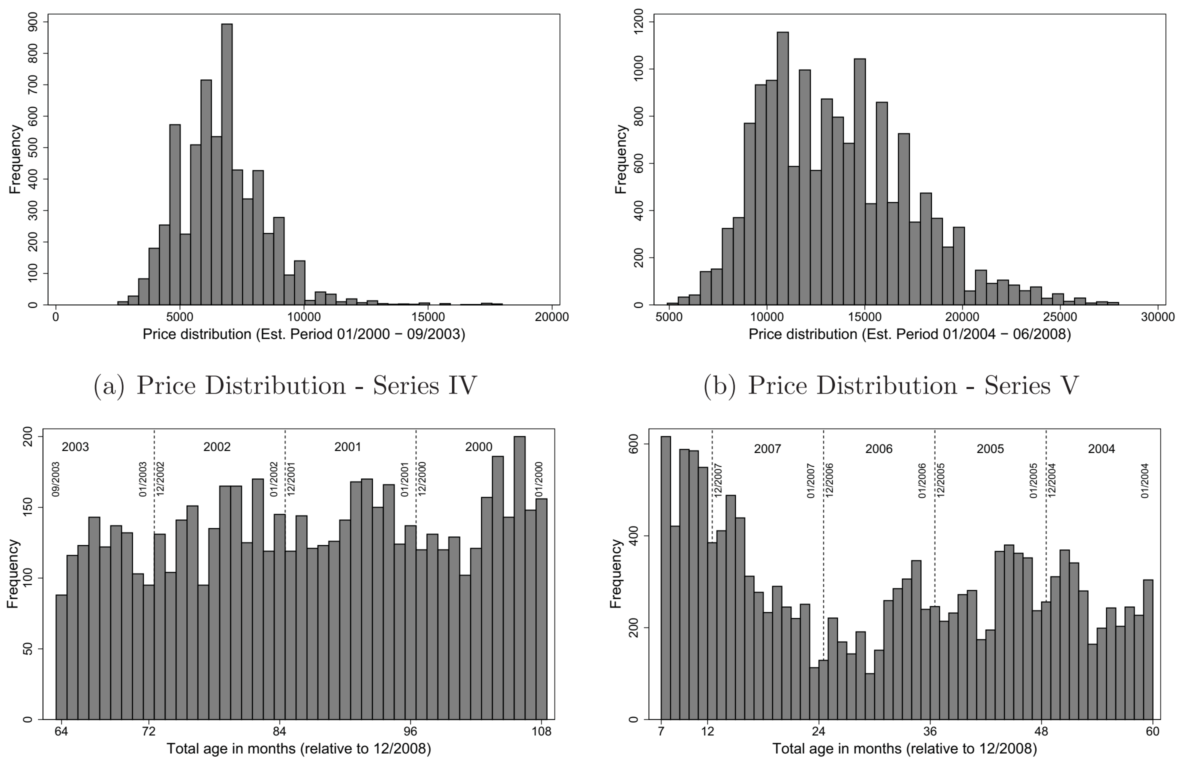
Returning to Table 4.2, observe that the average series IV (V) Golf has an age of 87 (30) months, exhibits an odometer reading of 115,978 km (55,343 km) and is offered at a price of €6,711 (€13,605), as shown in Panel A (B). In both series, most of the cars have five doors, a manual gearbox, and a metallic paint. However, the frequency of diesel cars, seat heatings, all-wheel drives, and cruise controls is considerably higher for the newer series V than for series IV. Also note that a large majority of offers originates from professional car dealers, as indicated by the dummy *private_seller* being equal to zero.

As we would expect, a correlation analysis for *price* yields a strong negative correlation coefficient with *totalage* ($\rho = -0.85$) and with *mileage* ($\rho = -0.78$). Conversely, *power* ($\rho = 0.45$), *diesel* ($\rho = 0.11$), *five-door* ($\rho = 0.16$), and all of the considered extras are significantly positively related to the price of a car.[14]

[14]Among the explanatory variables, we find that *totalage* and *mileage* co-move at a degree of $\rho = 0.77$. While in general collinearity among the explanatory variables can be problematic, our sample size is sufficiently large to produce precise parameter estimates.

While not listed in Table 4.2, another important determinant for the price of a car is its color. We therefore additionally include a set of color-dummies to control for their impact on price, where the effects are measured relative to *black*. We find that the prices are indeed somewhat responsive to different colors. For the sake of clarity, however, in the discussion below the respective coefficients for the color-dummies are not reported, but are available from the authors upon request.

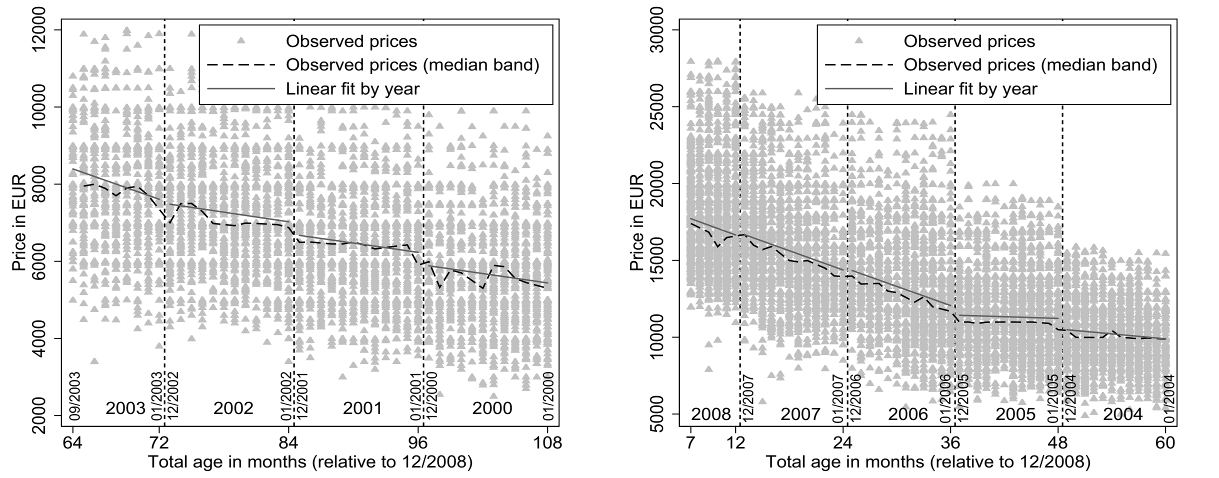
Figure 4.4: Distributions of Price and Age - VW Golf



(a) Price Distribution - Series IV
(b) Price Distribution - Series V
(c) Age Distribution - Series IV
(d) Age Distribution - Series V

Next, consider the distribution of price and age of the cars, which are depicted in Figure 4.4. For series IV, the prices are approximately normally distributed around the mean at €6,711 (Figure 4.4a). Likewise, the price pattern for series V concentrates around €13,605, though slightly more dispersed than the latter (Figure 4.4b). Regarding age, we find a very balanced distribution for series IV (Figure 4.4c), indicating that our sample contains a sufficient number of observations for each FR-date in the estimation period. The same applies for series V (Figure 4.4d), although the number of offers fluctuates considerably more across registration dates. The highest frequency of offers is observed for relatively new cars, i.e. around an age of 7 to 15 months relative to 12/2008, which is not surprising given the high number of professional car dealers that is active in this market segment (see Table 4.2).

Figure 4.5: Relation Between Price and Age - VW Golf



(a) Series IV

(b) Series V

A graphical inspection of the relation between price and total age of the cars in Figure 4.5 yields a first indication that information on the month of first registration may not be sufficiently utilized. First, consider the graph for series IV (Figure 4.5a): Although the linear fits imply that prices decline in age, just as we would expect, surprisingly much of this adjustment takes place *across* rather than *within* the vintages, as indicated by remarkable drops where the FR-year changes. For instance, for almost all first registration dates within the years 2002 and 2001 the respective median price level remains approximately constant, but between the years the median price slumps considerably downward. Stated differently, despite a car first registered in 12/2001 factually is almost a full year older than a car first registered in 01/2001, the median price does not change considerably during this period. In contrast to that, there is a pronounced price differential between 01/2001 and 12/2000 though ceteris paribus these cars merely differ by one marginal month of age. Similar discontinuities arise between the years 2003 and 2002, and also between 2002 and 2001. Likewise, also the price pattern for series V shown in Figure 4.5b exhibits distinct discontinuities at 12/2004, 12/2005. To a lesser extent, the same applies to 12/2006 and 12/2007, even though the precise age obviously has an increased impact on the prices of younger cars, as indicated by the steeper slopes of the linear fits within these vintages. Moreover, also for all series of the other models in our data we find qualitatively similar patterns (see Figures 4.8 - 4.10 in Appendix 4.5). Importantly, however, there is no evident rationale why a car from the same series should lose in value just because it displays a different figure in its registration year. However, all analyzed price patterns suggest that this is what actually happens.

4.3 Empirical Analysis

A crucial feature of complex, or composite goods like the used cars in our sample is that they can theoretically be decomposed into their constituent characteristics. Using a hedonic regression model, we are thus able to obtain estimates for the contributory value of each individual feature of a car, which in combination describe its overall quality and account for its aggregate value, or price. In light of the findings inferred from a mere inspection of the data, particularly the identification of the relation between the price and the age of the used cars is at the core of our interest. Before we present the results of the hedonic regression analysis, we briefly discuss the structural model our estimations are based upon. In a further step we control for potential pitfalls in our data, discuss an alternative specification, and test for the validity of our findings with respect to the models of other makes in our dataset.

4.3.1 Estimation Model and Predictions

Since our main interest is to identify to what extent the observed prices fully reflect the information provided on the age of a car, i.e. the precise date of its first registration, we need to separate the effects of the month and the year of age in the estimation. In doing so, we need to account for the fact that the patterns illustrated in Figure 4.5 indicate that the value adjustment due to a change in the FR-year is not constant across different years. In particular, we decompose the variable *fr_year* into dummies for each vintage, which thus capture the effect of an individual registration year on *price* if included as regressors. For example, the effect of vintage 2002 is measured by *year2002*, which takes value 1 if the first registration occurred during year 2002, and 0 otherwise. Since the estimation window for Golf series IV (V) comprehends cars from four (five) subsequent years, this corresponds to four (five) vintage-dummies, which by design are perfectly correlated. Thus, is suffices to include only three (four) of them in the regression on *price*. As we drop *year2003* (*year2008*) from the regression for series IV (V), the resulting coefficients for the included vintage-dummies are to be interpreted as the total price differential relative to the youngest cars in the relevant reference years, i.e. those registered in 12/2003 (12/2008).

As also implied by Figure 4.5, the slope of the price pattern - i.e. the impact of a marginal month - remains roughly constant within a vintage, but varies across vintages. Including the measure *totalage* in the estimation model is therefore not sufficient to produce precise estimates for the effect of a marginal month on *price*. Instead, we take a different tack. First, we define the variable

$$month \equiv 13 - fr\_month,$$

which reflects the age of a car in units of months within a particular vintage.[15] For instance, if a car was first registered in December (January), i.e. $fr\_month = 12$ ($fr\_month = 1$), we have that $month = 1$ ($month = 12$), which captures that these are the cars with the youngest (oldest) age in the respective vintage. Second, to account for the variation in the slopes across vintages, we interact this variable with each of the year dummies, and enter the resulting interaction terms into the regression on *price*. Thereby, we are able to estimate the impact of a marginal month separately for each vintage. For instance, the effect of one additional month of age in year 2002 is reflected through the coefficient of

$$month2002 \equiv month \times year2002,$$

where $month2002 \in [1, 12]$ for $year2002 = 1$, and 0 otherwise. For all other years, the interaction terms are defined and labeled accordingly. Formally, this specification corresponds to the assumption of a piecewise linear relationship between *price* and the precise age, which seems legitimate given the devolution of the price paths depicted above.[16] All other regressors are assumed to enter linearly into the regression model, which for the sample of Golf series IV is thus given by

$$\begin{aligned} price = \alpha &+ \beta_{m2000} \cdot month2000 + ... + \beta_{m2003} \cdot month2003 \\ &+ \beta_{y2000} \cdot year2000 + \beta_{y2001} \cdot year2001 + \beta_{y2002} \cdot year2002 \\ &+ \beta_{mileage} \cdot mileage + \beta_{power} \cdot power + \vec{\beta}\mathbf{X} + u, \end{aligned}$$

where $\mathbf{X}$ includes all quality controls and u is an error term. The model for series V is analogously defined, only that the vintage-dummies and interaction terms are adjusted to the respective estimation period.

[15] Note that this definition is equivalent to the second term on the RHS in the definition of *totalage*.

[16] Moreover, note that this specification is essentially equivalent to the one employed in Englmaier and Schmöller (2009b) to identify the "birthday effect" in the data of virtual football players. Since the main purpose of this paper is to test for the external validity of this effect, this resemblance is highly convenient.

This approach allows us to analyze the price pattern for potential discontinuities due to a change of the registration year, which reversely would imply that the information on the precise age - i.e. the month of first registration - is not efficiently utilized. Thereby, the underlying rationale is that each regression returns two alternative measures for the price differential between two vintages, because we separately account for the impact of years and months. First, from the coefficients for the vintage-dummies we are able to calculate the total difference in the values between two subsequent FR-years. In the estimation for series IV, for instance, the coefficient β_{y2002} reflects the price differential between a car registered in 12/2002 relative one registered in to 12/2003. Second, holding the vintage (and all other variables) constant the coefficients of the interaction terms provide us with an estimate of the impact of a marginal month of age within a particular year, e.g. β_{m2003} for 2003.

Clearly, if the prices of used cars decline continuously in their precise age in months, the value loss per year should be fully captured through the aggregated monthly effects within this timespan. Hence, any significant difference between the two measures identifies an additional value adjustment, i.e. a discontinuous drop in the price pattern, which can be attributed to the mere change of display in the FR-year. However, factually the quality of a car is completely unaffected by this event. More precisely, once we account for the impact of the finer information on the age of used cars in units of months by way of the interaction terms with *month*, the noisier information conveyed through the variable *fr_year*, i.e. the registration year, should be redundant and have no further impact on price. Translated to the estimation model, this is the case if and only if the coefficients on the vintage-dummies solely reflect the steady month-by-month decline of the price a car experiences during the course of a year. Returning to the above example, we would rationally expect that $12 \cdot \beta_{m2003} = \beta_{y2002}$. Formally, this implies the testable predictions regarding the relation between the coefficients of the vintage-dummies and those of the interaction terms as stated in Hypothesis 1.

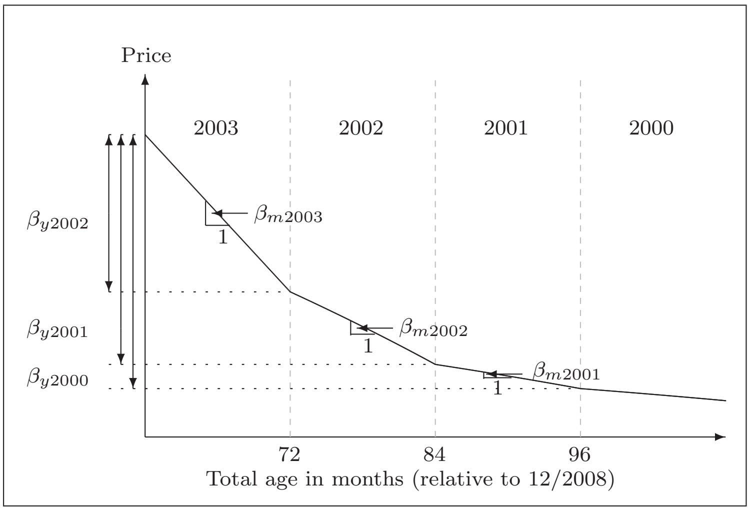
Hypothesis 1 *If the market value of used VW Golf cars declines continuously in their precise age in months, the value loss per year is fully captured through the aggregated marginal month-effects within this timespan. In the model framework, this is the case if and only if*

$$\beta_{y(t-1)} - \beta_{y(t)} = 12 \cdot \beta_{m(t)},$$

where $t \in [2000, 2003]$ *for series IV, and* $t \in [2004, 2008]$ *for series V.*

Since we measure the impact of the registration years relative to 2003 and 2008 for series IV and V, respectively, this immediately implies that $\beta_{y2003} = \beta_{y2008} = 0$. A graphical illustration of the estimation model is depicted in Figure 4.6, where we representatively focus on series IV. It shows how the prices should evolve under the above assumption of a piecewise linear relationship between price and age, and given that Hypothesis 1 holds.

Figure 4.6: Structural Model and its Predictions



4.3.2 Hedonic Regression Results

To test the validity of the predictions from Hypothesis 1, we start out with hedonic OLS regressions on *price*, where in addition to the age variables we include the full set of controls available in our data. The results from this approach are presented in Table 4.3 for VW Golf series IV and V, respectively.

As indicated by the values of R^2 and the F-statistics, in both regressions the underlying estimation model predicts a substantial share of the variability in the data. Intuitively, this reflects that much of the observed price dispersion can be attributed to variations in the set of included regressors, yielding a considerable degree of explanatory power. This intuition is further substantiated in Figure 4.7, which plots the predictions from the estimation model against the actual prices observed in the data, indicating a considerable goodness of fit.

Returning to Table 4.3, Panel A states the resulting coefficients and standard errors for the 6,034 observations in the subgroup of series IV Golfs. First, consider the set of control variables. Consistent with the predictions from the above correlation analysis, with exception

Table 4.3: Determinants of Price - VW Golf (OLS)

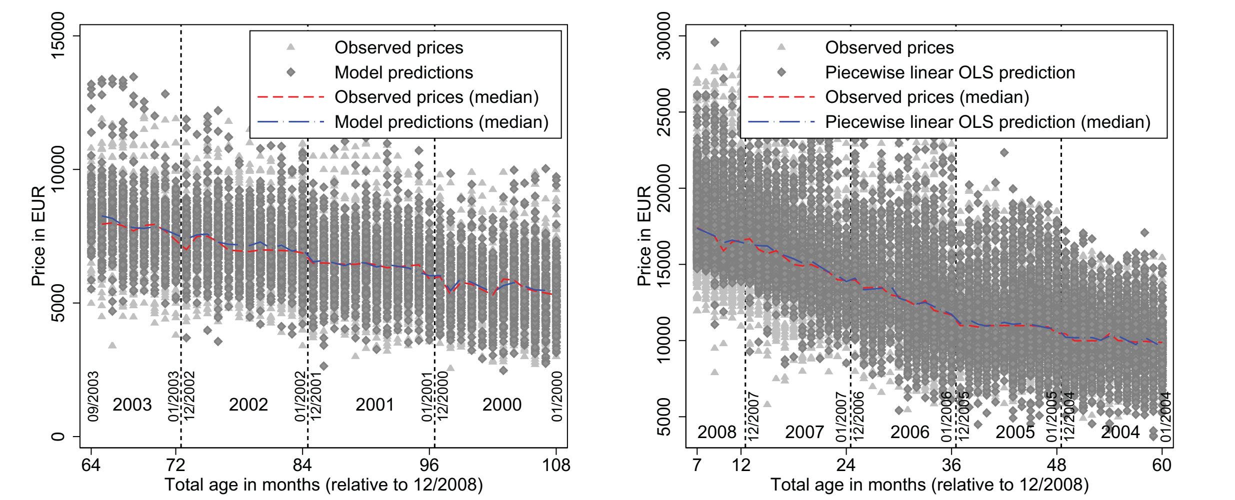
Panel A. Series IV (Est. Period: 01/2000 - 09/2003)			Panel B. Series V (Est. Period: 01/2004 - 06/2008)		
Variable	Coefficient	Std. Dev.	Variable	Coefficient	Std. Dev.
month2000	-25.93***	(6.71)	month2004	-15.42**	(7.09)
month2001	-20.13***	(7.12)	month2005	-5.09	(6.87)
month2002	-20.14***	(7.53)	month2006	-43.11***	(8.23)
month2003	-54.80***	(15.43)	month2007	-109.29***	(9.27)
			month2008	-119.73***	(19.74)
year2000	-1,734.41***	(141.40)	year2004	-4,770.43***	(194.00)
year2001	-1,326.71***	(142.90)	year2005	-4,312.65***	(194.66)
year2002	-816.96***	(142.91)	year2006	-3,357.15***	(197.47)
			year2007	-1,645.13***	(193.46)
mileage	-0.02***	(0.00)	mileage	-0.04***	(0.00)
power	37.02***	(1.39)	power	68.29***	(0.78)
diesel	547.03***	(35.14)	diesel	913.03***	(29.02)
five_door	290.88***	(30.20)	five-door	279.69***	(31.09)
auto_gearbox	-39.81	(56.47)	auto gearbox	748.67***	(45.40)
cruise_control	291.58***	(50.60)	cruise control	148.84***	(29.02)
seat heating	144.21***	(33.92)	seat heating	302.75***	(28.40)
all_wheel_drive	389.62***	(140.83)	all-wheel-drive	882.82***	(134.54)
sun_roof	17.28	(39.84)	sun roof	771.33***	(49.71)
leathertrim	830.26***	(133.93)	leathertrim	1,504.34***	(87.46)
metallic_paint	68.58**	(34.49)	metallic paint	47.50	(34.99)
private_seller	-220.73***	(30.87)	private seller	-365.69***	(45.68)
Intercept	7,402.98***	(157.19)	Intercept	12,483.68***	(201.58)
Color dummies	✓		Color dummies	✓	
R^2	0.66		R^2	0.84	
N	6,034		N	15,247	
$F_{(25;6,008)}$	385.59		$F_{(27;15,219)}$	2,483.39	

Notes: Dependent variable is *price*. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. The coefficients of the vintage-dummies are to be interpreted relative to the reference years 2003 (Panel A) and 2008 (Panel B), respectively. Robust standard errors are stated in parentheses.

of *auto-gearbox* and *sun-roof*, all considered features have a statistically significant impact on the price of a used car, indicating that the sellers take these factors into account when they choose their price. For instance, a marginal kilometer on the odometer reduces the price of a series IV Golf on average by an amount of €0.02. Conversely, as indicated by a positive sign of the respective coefficients, all else equal the average price of a car increases with the horsepower of the engine, if it runs on diesel rather than petrol, and if it has five instead of three doors, all significant at the 99% confidence level. Similarly, additional extras like a seat-heating or a leather interior are also significantly positively related to the price level. The results for series V Golfs in Panel B are qualitatively similar.

Second, in both panels also the coefficients for the vintage-dummies are of large magnitude and significant at the highest level, indicating that there is indeed a declining relationship between the age of an used VW Golf and its price. For example, all else equal, the fact that β_{y2002} is negative in Panel A indicates that the price for a car from series IV registered in

Figure 4.7: Model Fit - Predicted vs. Observed Prices



(a) Series IV (b) Series V

12/2002 is on average by €817 lower compared to one that was registered in 12/2003. As pointed out above, if *price* declines continuously in age, this significant reduction will solely reflect the aggregated effects of the value loss per month during the year 2003. In line with this argument, observe that the coefficient for *month2003*, i.e. β_{m2003}, has a negative sign and is statistically significant, indicating that a marginal month does have an impact on the price. The same applies for all other years in the estimation period for series IV.

However, a test of the theoretical predictions reveals surprising results. Aggregated over the whole period of 2003, the price reduction due to the month-per-month decline only amounts to $-€\,660 = 12 \cdot \beta_{m2003}$, or 80.7% of the total price differential of $-€\,817$ between 2003 and 2002, as implied by the coefficient β_{y2002}. Irrespective of the steady decline per month, due to the change in the FR-year a car from 2002 thus on average looses some additional 19.3%, or €157, in value. A Wald-Test confirms that this drop in the price is statistically significant at the 10%-level (p-value: 0.0696). In relation to the measured impact of one marginal month in 2003, i.e. $\beta_{m2003} = -€\,55$, and counterfactual to a fully continuous decline, the price pattern exhibits a substantial discontinuity at this point. Similarly, the total price differential between the years 2002 and 2001 is given by $\beta_{y2001} - \beta_{y2002} = -€\,510$, while the steady month-by-month decline during 2002 only amounts to $-€\,240 = 12 \cdot \beta_{m2002}$, or 47.1% of the former. The remaining 52.9% ($-€\,270$) establish another discontinuity between 01/2002 and 12/2001, which is statistically significant at the highest level (p-value: 0.0003). Finally, since the month-by-month reduction in price during 2001, i.e. $12 \cdot \beta_{m2001} = -€\,240$, accounts only for 59.0% of the total drop between 2001 and 2000, which is given by $\beta_{y2000} - \beta_{y2001} = -€\,407$, this identifies a third discontinuity of 41.0% ($-€\,167$) of the total decline located between these two vintages (p-value: 0.0136).

In Panel B, a similar pattern arises for VW Golf series V. Relative to 12/2008, the price for a car registered in 12/2007 is on average lower by €1,645. Since only 87.5% ($-€\,1440 = 12 \cdot \beta_{m2008}$) are explained by the aggregated month-by-month decline during 2008, 12.5% (€205) of the total price differential are caused by the mere change in the year count (p-value: 0.0175). Analogously, at 12/2006, 12/2005, and 12/2004, we find discontinuities of 23.6% (€404), 46.0% (€440), and 86.9% (€397), respectively, all statistically significant at the 1%-level.[17]

Table 4.4: Measured Discontinuities for the Different Models

Vintage	VW Golf Series IV	BMW 3 Series E46	Audi A4 Series B6	Opel Astra Series G
2000	-167**	-363***		-261*
	(41.0%)	(33.5%)		(87.9%)
2001	-270***	-233**	-608**	-172**
	(52.9%)	(39.3%)	(54.1%)	(37.4%)
2002	-157*	-460***	-389***	-154***
	(19.2%)	(39.0%)	(51.1%)	(23.4%)
2003		-167	-446***	
		(22.8%)	(31.2%)	
Vintage	**VW Golf Series V**	**BMW 3 Series E90**	**Audi A4 Series B7**	**Opel Astra Series H**
2004	-397***			-17
	(86.9%)			(2.5%)
2005	-440***	-766***	-328***	-286***
	(46.0%)	(46.7%)	(34.9%)	(79.9%)
2006	-404***	-1183***	-690***	-161*
	(23.6%)	(60.3%)	(30.7%)	(11.1%)
2007	-205**	-1546***		
	(12.5%)	(60.2%)		

Notes: All values are calculated from the corresponding regression coefficients. Asterisks denote statistical significance as indicated by a Wald-test at the 1%(***), 5%(**) or 10%(*) level. Shares of total drop are stated in parentheses.

By the same method, we also analyze the price patterns for the series of BMW 3, Audi A4, and Opel Astra within the defined estimation windows.[18] Table 4.4 shows an overview of the jumps in the price pattern that remain unexplained by the aggregated month-by-month decline within the preceding vintage, in absolute terms and as the relative share of the total value differential. Also for the other models, we find highly significant discontinuities of

[17]Observe that the impact of a marginal month in 2005 goes in the right direction, but is no longer statistically significant. Note that this implies that the prices do not adjust at all to the precise age within this year, making our result even stronger. In the calculation of the discontinuity at 12/2004, we account for β_{m2005} despite its insignificance.

[18]For the sake of clarity, the respective regression results are provided in Tables 4.9- 4.11 in Appendix 4.5.

sizeable magnitude at a change of the FR-year, indicating that this is a systematic pattern in this market and not idiosyncratic for the VW Golf. For instance, comparing the prices for BMW 3 series E90 registered in 01/2008 to those registered in 12/2007 we ceteris paribus find a dramatic value adjustment of €1,546, which accounts for 60.2% of the total price differential between these two registration dates. Importantly, except that the latter are one marginal month older than the former, the only real difference is that the last digit of the figure in the year count has changed.

Overall, we find considerable evidence that the price patterns for used cars exhibit distinct discontinuities at changes of the FR-year, indicating that their prices do not sufficiently react to their precise age as measured in units of months, which leads us to reject Hypothesis 1. Our intuition is that this effect arises because the sellers do not efficiently account for the stated month of first registration when forming their evaluation of a car, even though it represents much finer information than the year of first registration alone. As a consequence, the average price for used cars first registered in January of some vintage is significantly higher than that for the ones registered in December of the respective preceding vintage, even though they belong to the same series and their quality is held constant. Surprisingly, although there are considerable amounts of money at stake, the prices for used cars thus exhibit an analog to the “birthday effect”, which was documented in Englmaier and Schmöller (2009b) for the virtual *HT* economy. Stated differently, the findings we are able to document in our data establish suggestive evidence for the external validity of the “birthday effect”. Moreover, in light of the fact that a large fraction of the offers is made by professional car dealers, who should have considerable expertise, this finding is even more puzzling.

4.3.3 Robustness of Results

Before we turn to a discussion of possible explanations and the economic implications of our main finding, we briefly present a series of robustness tests. First, we provide a short summary of the results from alternative regression procedures that are used to account for potential pitfalls in the data. Subsequently, we analyze whether the effects persist if we relax our above assumption of a constant slope of the price curve within each vintage.[19]

[19] In doing so, we again representatively focus on the VW Golf samples. The corresponding results for the other models are qualitatively similar and available form the authors upon request.

4.3.3.1 Alternative Estimation Methods

Especially since our analysis is based on field data, in deriving our results we naturally control for potential pitfalls like multicollinearity and heteroscedasticity. Each of our subsamples contains sufficiently many observations, hence we do not find the first to be a problem. However, testing for a non-constant variance of the residuals, both a Breusch-Pagan and a White test indicate heteroscedasticity in the data. In all of the above OLS regressions, we thus account for possible correlations of the residuals across observations by applying Huber-White-Sandwich-estimators to produce robust standard errors. An alternative remedy to heteroscedasticity is to perform a log-linear transformation of the dependent variable. If we use the natural logarithm of *price* instead of the raw values as the dependent variable, all results from the standard OLS approach prove highly robust and carry through virtually unattenuated.[20] For instance, the analysis for series IV reveals that 26.5% (at 12/2002), 53.0% (at 12/2001), and 42.9% (at 12/2000) of the total price differential to the preceding vintage are not explained through the steady month-by-month decline, all statistically significant and remarkably close to the shares derived in the standard OLS procedure. The same holds true for series V, where we measure highly significant slumps of 18.3% (at 12/2007), 30.6% (at 12/2006), 52.4% (at 12/2005), and 88.2% (at 12/2004) of the total value differential.

In addition, we also perform a robust regression approach that uses an iteratively re-weighted least squares procedure to control for the possibility of influential outliers. By this method, each observation is assigned a weight $\omega \in [0, 1]$ with higher weights given to better behaved observations, where ω is iteratively determined and extremely deviant cases are excluded from the analysis. The regression results obtained from this approach, which qualitatively mirror those from the OLS estimation in Table 4.3, are shown in Table 4.13 in Appendix 4.5. A test of the predictions from Hypothesis 1 substantiates that our results are also not driven by influential outliers.[21] Hence, within the specified estimation model, we conclude that the documented frictions prove to be robust across alternative estimation methods.

[20]See Table 4.12 in Appendix 4.5 for the regression output from this approach.

[21]For series IV, the measured discontinuities are given by 35.9% at 12/2002, 54.2% at 12/2001, and 33.7% at 12/2000, respectively. While we find no significant effect at 12/2007 for series V, all other discontinuities are significant on the highest level and given by 20.0% at 12/2006, 47.4% at 12/2005, and 90.6% at 12/2005.

4.3.3.2 Alternative model specification

Motivated by the insights from the graphical inspection of the data, the results presented so far rely on the assumption of a piecewise linear relationship, i.e. that the impact of a marginal month remains constant during the course of a year. In the following we relax this assumption and estimate a different model, in which we determine the impact on price separately for each individual month. In doing so, we first introduce the dummies *Jan* to *Dec* to indicate the month of the year. For instance, *Nov* takes value 1 if a car was first registered in November, and 0 otherwise. As a second step, we interact the latter with the year dummies, thereby obtaining a dummy variable for each possible registration date within the estimation period. For example,

$$\mathit{Nov2002} \equiv \mathit{Nov} \times \mathit{year2002},$$

such that *Nov2002* takes value 1 for cars first registered in 11/2002, and 0 otherwise. For all other FR-dates, an analogous variable is defined and labeled accordingly. Included as regressors, the coefficient for each of these variables measures the impact of a particular marginal month on the price, relative to the "youngest" FR-date in each subsample, i.e. 09/2003 for series IV and 06/2008 in series V. Intuitively, if the observed discontinuities are indeed caused by the change of the registration year, we would expect that the coefficients for two consecutive months *within* the same FR-year should not differ much, and most of the price adjustment should be located between December and January of the next vintage. The results from this approach are shown in Table 4.5. For simplicity, we focus the presentation on a four-month period around each change of the registration year.

First, consider the results for series IV in Panel A. The coefficients for *Feb2003* and *Jan2003* are of similar magnitude, and similarly for *Dec2002* and *Nov2002*. As confirmed by a t-test, in neither case the null hypothesis that the coefficients are equal can be rejected at p-values of 0.8133 and 0.7176, respectively. However, consistent with an over-proportional price adjustment due to a change of the registration year, $\beta_{Dec2003} - \beta_{Jan2002} < 0$ proves to be statistically significant at a p-value of 0.0691. Similarly, while neither the difference between the coefficients for *Feb2002* and *Jan2002* (p-value: 0.8349), nor that between *Nov2001* and *Dec2001* (p-value: 0.2101) are significantly different from zero, we find that $\beta_{Dec2002} - \beta_{Jan2001} < 0$ (p-value: 0.0023). Yet, we do not find a similar pattern between 2001 and 2000.

Table 4.5: Determinants of Price - VW Golf Monthwise Model (OLS)

Panel A. Series IV (Est. Period: 01/2000 - 09/2003)			Panel B. Series V (Est. Period: 01/2004 - 06/2008)		
Variable	Coefficient	Std. Dev.	Variable	Coefficient	Std. Dev.
Nov 2000	-1,609.24***	(163.24)	Nov 2004	-4,250.5***	(102.37)
Dec 2000	-1,398.07***	(167.03)	Dec 2004	-4,099.9***	(101.91)
Jan 2001	-1,462.98***	(164.24)	Jan 2005	-3,943.9***	(111.02)
Feb 2001	-1,245.52***	(172.18)	Feb 2005	-3,772.7***	(113.09)
Nov 2001	-1,100.60***	(169.05)	Nov 2005	-3,771.0***	(118.92)
Dec 2001	-1,263.69***	(175.10)	Dec 2005	-3,762.9***	(110.15)
Jan 2002	-852.58***	(173.70)	Jan 2006	-3,254.3***	(112.67)
Feb 2002	-879.01***	(169.47)	Feb 2006	-3,322.3***	(108.36)
Nov 2002	-636.23***	(175.62)	Nov 2006	-2,699.0***	(135.49)
Dec 2002	-685.00***	(172.15)	Dec 2006	-3,005.0***	(126.18)
Jan 2003	-410.14**	(188.43)	Jan 2007	-2,136.2***	(156.17)
Feb 2003	-447.32**	(178.45)	Feb 2007	-2,208.9***	(162.89)
			Nov 2007	-1,275.8***	(104.65)
			Dec 2007	-879.54***	(113.45)
			Jan 2008	-379.84***	(124.09)
			Feb 2008	-836.11***	(102.39)
mileage	-0.02***	(0.00)	mileage	-0.04***	(0.00)
power	36.98***	(1.40)	power	68.24***	(0.78)
diesel	541.75***	(35.06)	diesel	900.56***	(29.13)
five-door	289.09***	(30.28)	five-door	286.98***	(31.31)
auto gearbox	-34.55	(56.82)	auto gearbox	741.60***	(45.40)
private seller	-221.60***	(30.99)	private seller	-360.70***	(45.80)
Intercept	7,179.21***	(170.88)	Intercept	11,886.07***	(101.44)
Extras	✓		Extras	✓	
Color dummies	✓		Color dummies	✓	
All other months	✓		All other months	✓	
R^2	0.66		R^2	0.84	
N	6,034		N	15,247	
$F_{(59;5,971)}$	160.75		$F_{(71;15,175)}$	960.31	

Notes: Dependent variable is *price*. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. The coefficients of the vintage-dummies are to be interpreted relative to the reference months 09/2003 (Panel A) and 06/2008 (Panel B), respectively. Robust standard errors are stated in parentheses. To save on space, the details on the extras are omitted from the presentation. All results from the standard OLS regressions fully carry over.

Next, turn to Panel B, which states the coefficients obtained for series V. For every turn of the registration year in the estimation period, we find a clear trend for the price differential to increase, with pronounced effects between January and December. Though all coefficients around the change from 2008 to 2007 turn out to differ significantly, in line with our argument, the largest adjustment is located between 12/2008 and 01/2007. In addition, between the years 2007 and 2006 only the difference of the coefficients for *Jan2007* and *Dec2006* is statistically significant at the highest level (p-value: 0.0000). The same holds true around the change from 2006 to 2005, where $\beta_{Dec2005} - \beta_{Jan2006} < 0$ at a p-value of 0.0000. Between the years 2005 and 2004, however, again none of the displayed coefficients differ significantly.

Importantly, by design of this approach, the estimates for the impact of an individual month are based on rather few observations within a relatively short time interval, implying a considerable degree of noise due to price fluctuations in the data. Despite this fact, we still find clear evidence for a significant downward adjustment of the price at a turn of the registration year. Moreover, an analysis of the other models yields similar results. Hence, we take this as further indication that it is indeed the change of the registration year that triggers the discontinuities in the price pattern.

4.4 Discussion and Conclusion

We employ a hedonic regression analysis to examine empirically to what extent the stated prices for used cars reflect available information. Based on detailed field data on used car offers from the online vehicle market platform *mobile.de*, we find strong evidence for biased information processing. Despite the large monetary stakes involved, our findings suggest that people in this market systematically fail to efficiently aggregate the information provided on specific attributes of the items on sale. In particular, although the precise date of first registration is clearly stated, the pattern of observed prices exhibits sizeable discontinuities, indicating that a substantial fraction of the value adjustment due to the age of a car is located where the FR-year changes. As a consequence, across two consecutive vintages the price differential for cars with otherwise close-by registration dates is significantly larger than rationally justified, given that they only marginally differ in their precise age. This finding proves highly robust across several estimation approaches and if we separately control for the impact of each possible registration date, and indicates that an evaluation bias like the “birthday” effect can have economic implications in the real world.

The fact that we are able to provide suggestive evidence for a systematic friction in an otherwise highly competitive market, where in addition individual choices are conceivably subject to profound deliberations, naturally raises two closely related questions. First, what are the driving forces behind this effect? And second, what are the economic consequences of this finding? Along the lines of Englmaier and Schmöller (2009b) and Englmaier and Schmöller (2009a),[22] where we study both the behavior of buyers and sellers in the virtual *HT* economy, there are several possible answers to these questions.

[22] See Chapters 2 and 3 of this dissertation, respectively.

First, recall that our findings derive from asking prices stated by the sellers of used cars. Therefore, one possibility is that they act strategically and deliberately charge higher prices for cars that are factually not much older than others, but just belong to the next higher vintage in terms of their registration year. Intuitively, this would be a rational reaction if the population of potential buyers at least partially consists of agents who are inattentive to the precise age of the offered cars. Stated differently, what we observe could be the result of the sellers trying to exploit a bias on the demand side of the market for used cars, and it is actually the buyers who inefficiently utilize the provided information conveyed through the FR-month. Alternatively, it may be the sellers themselves, who disproportionately cling to the figure displayed in the FR-year, while in turn under-weighing the information embodied in the FR-month, such that the latter will not be efficiently incorporated into their pricing considerations. Finally, the documented behavior may apply to both sides of the market. Although in this environment, other than in the auction market analyzed in Englmaier and Schmöller (2009b), the economic consequences are likely to be mitigated by the possibility of negotiation, from the perspective of rational buyers a substantial fraction of cars will be overpriced, potentially leading to too little trade.

Several extensions to this research suggest themselves. As shown in Englmaier and Schmöller (2009b), a potential source for this effect may be linked to the design of the filter mechanism, which the people can use to screen and cross compare different offers. Due to the fact that it is not possible to directly filter for the FR-month on platforms like *mobile.de*, it may be tempting to perceive this information as unimportant and to overly focus one's attention on the more salient FR-year. It would therefore be interesting to see whether the size of the discontinuities is affected by including this feature in the filter mechanism. Moreover, though we consider a large number of factors that have an significant impact on the price of a used car, clearly we are not able to include all possible dimensions, and the results may potentially affected by other confounding factors. Hence, controlled laboratory experiments could be a promising route in the study of the utilization of information. Although field evidence might have higher external validity than controlled laboratory results, the latter allow for systematic variations of the market design features and the degree of valuation uncertainty, while at the same time the asking and final prices can be observed and detailed information on the decisions of each individual participant are provided.

In his seminal contribution to information economics, Akerlof (1970) employs the information asymmetries between buyers and sellers of used cars as his prime example to illustrate the famous "lemons-problem". Although adverse selection due to asymmetric information with respect to unobservables is undeniably still a major problem within this market, our findings suggest that inefficiencies may also arise with respect to observable characteristics. People seem to be inattentive to subtle, but nevertheless valuable details of the *available* information. Clearly, more research is needed to identify the driving forces and behavioral motivations behind an inefficient utilization of information, as we are able to document in the evaluation of complex goods.

4.5 Appendix

Table 4.6: Summary Statistics - BMW 3*

	Panel A. Series E46 (Est. Period: 01/2000 - 11/2004)				Panel B. Series E90 (Est. Period: 09/2005 - 08/2008)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	7,458	10,562	3,990	20,000	10,866	22,607	8,500	48,578
mileage [km]	7,458	113,130	6,163	227,000	10,866	58,220	1,000	226150
power [kW]	7,458	110	75	185	10,866	126	85	247
totalage [months]	7,458	77	50	108	10,866	26	5	40
Dummies	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	4,373	58.6	58.6	0	3,007	27.7	27.7
(0 = petrol, 1 = diesel)	1	3,085	41.4	100.0	1	7,859	72.3	100.0
auto gearbox	0	5,730	76.8	76.8	0	7,579	69.8	69.8
(0 = manu., 1 = auto)	1	1,728	23.2	100.0	1	3,287	30.3	100.0
cruise control	0	5,067	67.9	67.9	0	4,453	41.0	41.0
(0 = no, 1 = yes)	1	2,391	32.1	100.0	1	6,413	59.0	100.0
seat heating	0	3,793	50.9	50.9	0	3,309	30.5	30.5
(0 = no, 1 = yes)	1	3,665	49.1	100.0	1	7,557	69.6	100.0
all-wheel-drive	0	7,222	96.8	96.8	0	10,414	95.8	95.8
(0 = no, 1 = yes)	1	236	3.2	100.0	1	452	4.2	100.0
sun roof	0	4,735	63.5	63.5	0	6,848	63.0	63.0
(0 = no, 1 = yes)	1	2,723	36.5	100.0	1	4,018	37.0	100.0
leathertrim	0	5,359	71.9	71.9	0	7,684	70.7	70.7
(0 = no, 1 = yes)	1	2,099	28.1	100.0	1	3,182	29.3	100.0
metallic paint	0	1,091	14.6	14.6	0	2,957	27.2	27.2
(0 = no, 1 = yes)	1	6,367	85.4	100.0	1	7,909	72.8	100.0
private seller	0	5,544	74.3	74.3	0	10,350	95.3	95.3
(0 = no, 1 = yes)	1	1,914	25.7	100.0	1	516	4.8	100.0

* This sample was the first we collected. At that time, we exclusively considered five-door cars to control for convertibles, which could not be distinguished from other three-door vehicles. In the other samples an upgrade of the parsing software allowed us filter directly for convertibles.

Table 4.7: Summary Statistics - Audi A4*

	Panel A. Series B6 (Est. Period: 10/2001 - 10/2004)				Panel B. Series B7 (Est. Period: 04/2005 - 09/2007)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	4,055	11,986	4,000	21,900	6,938	18,249	6,900	39,990
mileage [km]	4,055	111,448	9,729	233,550	6,938	90,143	1,000	234000
power [kW]	4,055	106	66	169	6,938	115	66	294
totalage [months]	4,055	68	51	87	6,938	34	16	45
Dummies	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	2,655	48.1	48.1	0	1,440	20.8	20.8
(0 = petrol, 1 = diesel)	1	2,864	51.9	100.0	1	5,498	79.2	100
auto gearbox	0	3,936	71.3	71.3	0	4,735	68.3	68.3
(0 = manu., 1 = auto)	1	1,583	28.7	100.0	1	2,203	31.8	100
cruise control	0	3,877	70.3	70.3	0	3,236	46.6	46.6
(0 = no, 1 = yes)	1	1,642	29.8	100.0	1	3,702	53.4	100
seat heating	0	2,641	47.9	47.9	0	2,126	30.6	30.6
(0 = no, 1 = yes)	1	2,878	52.2	100.0	1	4,812	69.4	100
all-wheel-drive	0	4,892	88.6	88.6	0	5,752	82.9	82.9
(0 = no, 1 = yes)	1	627	11.4	100.0	1	1,186	17.1	100
sun roof	0	4,719	85.5	85.5	0	5,974	86.1	86.1
(0 = no, 1 = yes)	1	800	14.5	100.0	1	964	13.9	100
leathertrim	0	4,545	82.4	82.4	0	5,337	76.9	76.9
(0 = no, 1 = yes)	1	974	17.7	100.0	1	1,601	23.1	100
metallic paint	0	938	17.0	17.0	0	1,710	24.7	24.7
(0 = no, 1 = yes)	1	4,581	83.0	100.0	1	5,228	75.4	100
private seller	0	4,276	77.5	77.5	0	6,506	93.8	93.8
(0 = no, 1 = yes)	1	1,243	22.5	100.0	1	432	6.2	100

* The Audi A4 production line includes no three-door version.

Table 4.8: Summary Statistics - Opel Astra*

	Panel A. Series G (Est. Period: 01/2000 - 12/2003)				Panel B. Series H (Est. Period: 05/2004 - 10/2007)			
Variable	Obs.	Mean	Min	Max	Obs.	Mean	Min	Max
price [EUR]	2,825	5,332	1,750	9,950	8,358	10,820	3,900	19,900
mileage [km]	2,825	108,788	2,732	206,000	8,358	63,969	1,000	206,000
power [kW]	2,825	69	48	188	8,358	80	55	177
totalage [months]	2,825	85	61	108	8,358	35	15	56
Dummies	Value	Frequency	Percent	Cum.	Value	Frequency	Percent	Cum.
diesel	0	2,104	74.5	74.5	0	3,817	45.7	45.7
(0 = petrol, 1 = diesel)	1	721	25.5	100.0	1	4,541	54.3	100.0
five-door	0	463	16.4	16.4	0	760	9.1	9.1
(0 = three, 1 = five)	1	2,362	83.6	100.0	1	7,598	90.9	100.0
auto gearbox	0	2,574	91.1	91.1	0	7,878	94.3	94.3
(0 = manu., 1 = auto)	1	251	8.9	100.0	1	480	5.7	100.0
cruise control	0	2,649	93.8	93.8	0	1,007	12.1	12.1
(0 = no, 1 = yes)	1	176	6.2	100.0	1	7,351	88.0	100.0
seat heating	0	2,653	93.9	93.9	0	7,472	89.4	89.4
(0 = no, 1 = yes)	1	172	6.1	100.0	1	886	10.6	100.0
sun roof	0	2,679	94.8	94.8	0	8,202	98.1	98.1
(0 = no, 1 = yes)	1	146	5.2	100.0	1	156	1.9	100.0
leathertrim	0	2,738	96.9	96.9	0	7,968	95.3	95.3
(0 = no, 1 = yes)	1	87	3.1	100.0	1	390	4.7	100.0
metallic paint	0	520	18.4	18.4	0	1,390	16.6	16.6
(0 = no, 1 = yes)	1	2,305	81.6	100.0	1	6,968	83.4	100.0
private seller	0	2,395	84.8	84.8	0	7,832	93.7	93.7
(0 = no, 1 = yes)	1	430	15.2	100.0	1	526	6.3	100.0

* None of the observations for Opel Astra included an all-wheel drive.

Table 4.9: Determinants of Price - BMW 3 (OLS)

Panel A. Series E46 (Est. Period: 01/2000 - 11/2004)			Panel B. Series E90 (Est. Period: 09/2005 - 08/2008)		
Variable	Coefficient	Std. Dev.	Variable	Coefficient	Std. Dev.
month2000	8.25	(-10.81)	month2005	-14.60	(-48.07)
month2001	-60.27***	(-10.95)	month2006	-73.19***	(-11.93)
month2002	-29.89***	(-10.87)	month2007	-65.31***	(-18.66)
month2003	-60.00***	(-10.05)	month2008	-84.93***	(-31.96)
month2004	-47.07***	(-13.33)			
year2000	-3,587.09***	(-132.14)	year2005	-6,171.47***	(-309.38)
year2001	-2,504.03***	(-133.36)	year2006	-4,529.36***	(-296.59)
year2002	-1,911.20***	(-129.86)	year2007	-2,565.81***	(-306.27)
year2003	-730.90***	(-126.73)			
mileage	-0.03***	(0.00)	mileage	-0.06***	(0.00)
power	41.97***	(-0.95)	power	81.02***	(-1.29)
diesel	573.46***	(-41.52)	diesel	1,681.84***	(-59.87)
auto gearbox	-56.35	(-44.06)	auto gearbox	1,087.39***	(-61.58)
private seller	38.44	(-42.45)	private seller	-475.41***	(-139.67)
Intercept	11,068.34***	(-143.93)	Intercept	16,439.34***	(-320.66)
Extras	✓		Extras	✓	
Color dummies	✓		Color dummies	✓	
R^2	0.74		R^2	0.81	
N	7,458		N	10,866	
$F_{(26;7,431)}$	792.58		$F_{(24;10,841)}$	1,462.92	

Notes: The coefficients of the vintage-dummies are to be interpreted relative to the reference years 2004 (Panel A) and 2008 (Panel B), respectively. Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. Since all coefficients for the extra are qualitatively similar to the ones obtained for VW Golf, they are omitted from the presentation to save on space.

Table 4.10: Determinants of Price - Audi A4 (OLS)

Panel A. Series B6 (Est. Period: 10/2001 - 10/2004)			Panel B. Series B7 (Est. Period: 04/2005 - 09/2007)		
Variable	Coefficient	Std. Dev.	Variable	Coefficient	Std. Dev.
month2001	96.60	(111.09)	month2005	-89.80***	(13.70)
month2002	-42.65***	(12.48)	month2006	-51.22***	(13.00)
month2003	-30.50**	(13.87)	month2007	-129.70***	(40.67)
month2004	-82.20***	(16.84)			
year2001	-3,315.33***	(282.24)	year2005	-3,190.15***	(371.25)
year2002	-2,190.99***	(164.94)	year2006	-2,249.79***	(376.30)
year2003	-1,430.03***	(172.72)			
mileage	-0.04***	(0.00)	mileage	-0.05***	(0.00)
power	26.73***	(1.68)	power	50.29***	(1.62)
diesel	808.96***	(56.31)	diesel	1,841.56***	(69.37)
auto gearbox	81.98	(59.44)	auto gearbox	461.37***	(57.51)
private seller	-128.01**	(62.26)	private seller	-333.42***	(121.69)
Intercept	14,308.04***	(226.10)	Intercept	17,570.93***	(413.66)
Extras	✓		Extras	✓	
Color dummies	✓		Color dummies	✓	
R^2	0.62		R^2	0.72	
N	4,055		N	6,938	
$F_{(25;4,029)}$	260.17		$F_{(23;6,914)}$	659.58	

Notes: The coefficients of the vintage-dummies are to be interpreted relative to the reference years 2004 (Panel A) and 2007 (Panel B), respectively. Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. Since all coefficients are qualitatively similar to the ones obtained for VW Golf, the details on the extras are omitted from the presentation to save on space.

Table 4.11: Determinants of Price - Opel Astra (OLS)

Panel A. Series G (Est. Period: 01/2000 - 12/2003)			Panel B. Series H (Est. Period: 05/2004 - 10/2007)		
Variable	Coefficient	Std. Dev.	Variable	Coefficient	Std. Dev.
month2000	-13.72*	(8.05)	month2004	-74.71***	(14.17)
month2001	-3.37	(8.27)	month2005	-55.80***	(6.81)
month2002	-23.89***	(9.14)	month2006	-5.56	(7.67)
month2003	-42.04***	(10.64)	month2007	-107.51***	(15.64)
year2000	-1,415.08***	(109.07)	year2004	-2,503.53***	(148.15)
year2001	-1,117.58***	(108.59)	year2005	-1,815.47***	(138.66)
year2002	-658.21***	(110.67)	year2006	-1,457.06***	(145.45)
mileage	-0.02***	(0.00)	mileage	-0.04***	(0.00)
power	17.30***	(1.57)	power	51.51***	(1.20)
diesel	123.01***	(43.27)	diesel	660.75***	(34.38)
five-door	328.06***	(40.68)	five-door	-402.37***	(53.27)
auto gearbox	80.40	(55.22)	auto gearbox	298.40***	(66.54)
private seller	-192.50***	(43.20)	private seller	-516.91***	(64.62)
Intercept	6,949.84***	(153.53)	Intercept	10,743.14***	(179.54)
Extras	✓		Extras	✓	
Color dummies	✓		Color dummies	✓	
R^2	0.61		R^2	0.65	
N	2,825		N	8,358	
$F_{(24;2,800)}$	203.91		$F_{(25;8,332)}$	597.13	

Notes: The coefficients of the vintage-dummies are to be interpreted relative to the reference years 2003 (Panel A) and 2007 (Panel B), respectively. Robust standard errors are stated in parentheses. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. Since all coefficients are qualitatively similar to the ones obtained for VW Golf, the details on the extras are omitted from the presentation to save on space.

Table 4.12: Determinants of Log-Price - VW Golf (OLS)

Panel A. Series IV (Est. Period: 01/2000 - 09/2003)			Panel B. Series V (Est. Period: 01/2004 - 06/2008)		
Variable	Coefficient	Std. Dev.	Variable	Coefficient	Std. Dev.
month2000	-0.0050***	(0.0011)	month2004	-0.0021***	(0.0006)
month2001	-0.0033***	(0.0011)	month2005	-0.0004	(0.0006)
month2002	-0.0028***	(0.0010)	month2006	-0.0028***	(0.0006)
month2003	-0.0048***	(0.0017)	month2007	-0.0054***	(0.0006)
			month2008	-0.0057***	(0.0011)
year2000	-0.2200***	(0.0170)	year2004	-0.2921***	(0.0111)
year2001	-0.1505***	(0.0168)	year2005	-0.2470***	(0.0111)
year2002	-0.0781***	(0.0162)	year2006	-0.1777***	(0.0111)
			year2007	-0.0843***	(0.0105)
mileage	-0.0000***	(0.0000)	mileage	-0.0000***	(0.0000)
power	0.0049***	(0.0002)	power	0.0047***	(0.0001)
diesel	0.0913***	(0.0050)	diesel	0.0700***	(0.0020)
five-door	0.0551***	(0.0044)	five-door	0.0226***	(0.0023)
auto gearbox	-0.0091	(0.0075)	auto gearbox	0.0462***	(0.0029)
private seller	-0.0297***	(0.0045)	private seller	-0.0313***	(0.0034)
Intercept	8.8972***	(0.0196)	Intercept	9.3941***	(0.0113)
Extras	✓		Extras	✓	
Color dummies	✓		Color dummies	✓	
R^2	0.67		R^2	0.85	
N	6,034		N	15,247	
$F_{(25;6,008)}$	459.34		$F_{(27;15,219)}$	2,871.86	

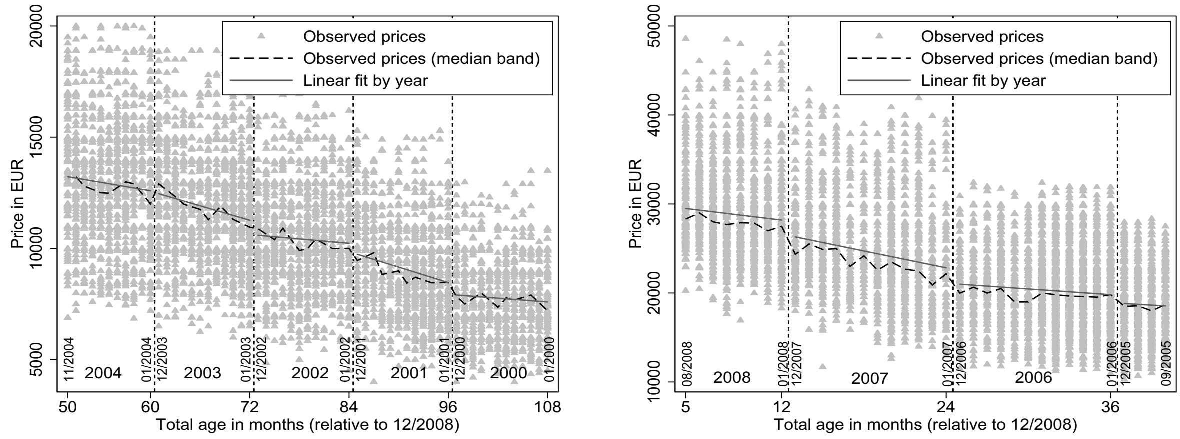
Notes: Dependent variable is *lnprice*. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. The coefficients of the vintage-dummies are to be interpreted relative to the reference years 2003 (Panel A) and 2008 (Panel B), respectively. Robust standard errors are stated in parentheses. To save on space, the details on the extras are omitted from the presentation. All results from the standard OLS regressions fully carry over.

Table 4.13: Determinants of Price - VW Golf (Robust Regression)

Panel A. Series IV (Est. Period: 01/2000 - 09/2003)			Panel B. Series V (Est. Period: 01/2004 - 06/2008)		
Variable	Coefficient	Std. Dev.	Variable	Coefficient	Std. Dev.
month2000	-24.18***	(6.57)	month2004	-17.69**	(7.77)
month2001	-21.96***	(6.69)	month2005	-3.45	(7.46)
month2002	-17.04**	(6.72)	month2006	-41.28***	(8.17)
month2003	-30.18**	(11.74)	month2007	-112.55***	(7.58)
			month2008	-157.86***	(15.30)
year2000	-1,408.55***	(111.00)	year2004	-4,953.22***	(157.96)
year2001	-1,011.25***	(110.25)	year2005	-4,511.53***	(159.30)
year2002	-564.79***	(110.16)	year2006	-3,570.59***	(160.09)
			year2007	-1,883.32***	(152.81)
mileage	-0.02***	(0.00)	mileage	-0.04***	(0.00)
power	30.39***	(0.84)	power	67.73***	(0.65)
diesel	590.29***	(29.00)	diesel	795.91***	(26.53)
five-door	370.28***	(26.96)	five-door	253.95***	(30.91)
auto gearbox	-19.06	(43.61)	auto gearbox	710.80***	(37.05)
private seller	-278.52***	(27.51)	private seller	-334.36***	(41.03)
Intercept	7,518.74***	(122.10)	Intercept	12,661.52***	(160.93)
Extras	✓		Extras	✓	
Color dummies	✓		Color dummies	✓	
R^2	0.68		R^2	0.85	
N	6,034		N	15,247	
$F_{(25,6008)}$	508.15		$F_{(27,15219)}$	3,232.22	

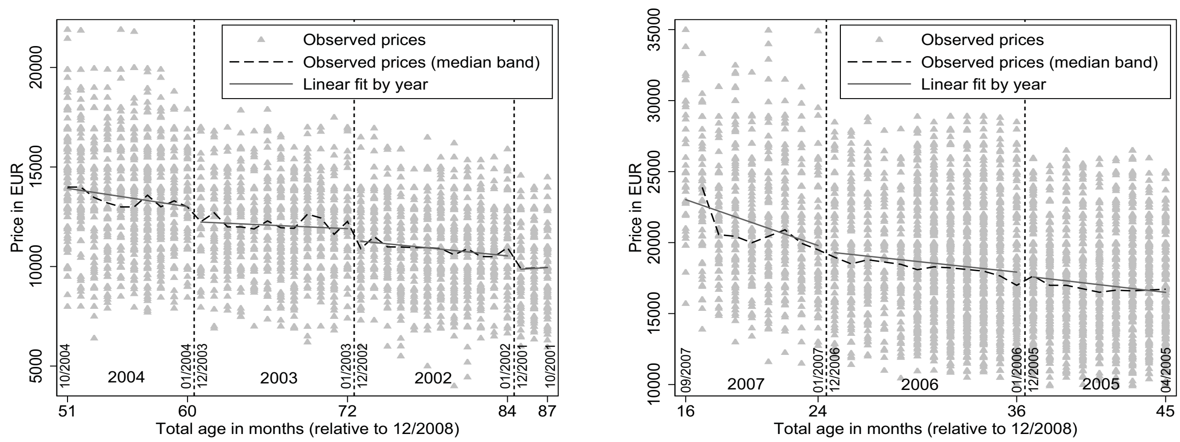
Notes: Dependent variable is *price*. The RREG procedure controls for influential outliers by computing point-specific weights for the contribution of each observation to the final regression. Asterisks denote statistical significance at the 1%(***), 5%(**) or 10%(*) level. The coefficients of the vintage-dummies are to be interpreted relative to the reference years 2003 (Panel A) and 2008 (Panel B), respectively. Standard errors are stated in parentheses. To save on space, the details on the extras are omitted from the presentation. All results from the standard OLS regressions fully carry over.

Figure 4.8: Relation Between Price and Age - BMW 3



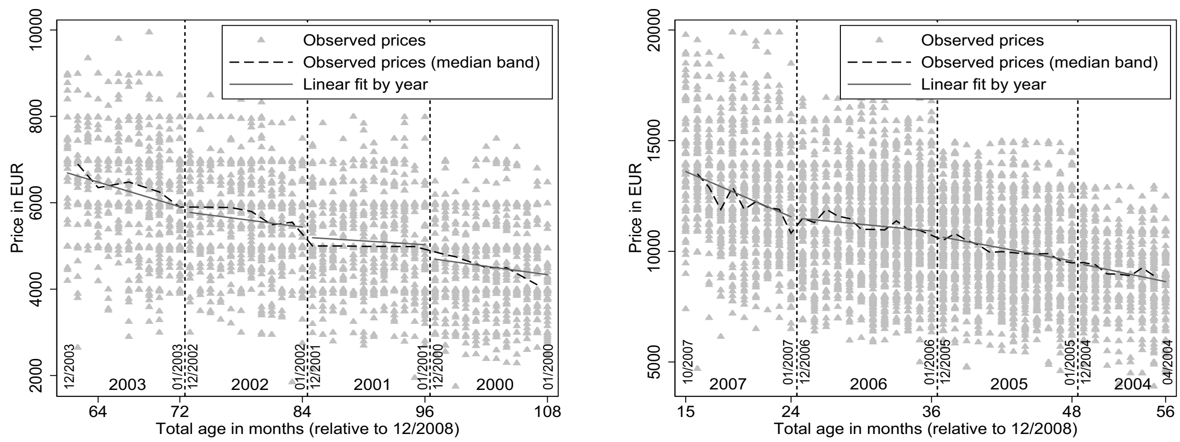
(a) Series E46 (b) Series E90

Figure 4.9: Relation Between Price and Age - Audi A4



(a) Series B6 (b) Series B7

Figure 4.10: Relation Between Price and Age - Opel Astra



(a) Series G (b) Series H

Chapter 5

Strategic Seller Actions in Auctions with Asymmetric Bidders

5.1 Introduction

A fundamental principle of an auction is to endogenously generate a price for an item, whenever the seller, or "auctioneer", is uncertain or has only incomplete information about the exact valuations of the potential buyers in the population. The final price the winner has to pay depends on various factors, like the auction mechanism implemented, the number of bidders, and the intensity of competition among them to name only a few. In economic theory, auctions are modeled as non-cooperative games of incomplete information and a very common assumption is that the buyers are ex ante symmetric with respect to the distribution of their valuations. While there exists an extensive body of literature for the case of symmetric bidders, far less general properties of asymmetric (first-price) auctions are identified so far. Among the rather few general properties, Maskin and Riley (2000) demonstrate that even one of the most principal findings in auction theory, the revenue equivalence theorem (Vickrey, 1961), is no longer generally valid when bidders are asymmetric.[1] However, Cantillon (2008) shows in a very general framework that in both first and second price auctions a reduction in the degree of asymmetry leads to an increase of expected revenue for the auctioneer. Intuitively, asymmetry hurts the seller, as competition among bidders is reduced resulting in a lower expected final price. We use this finding to analyze

[1] As a consequence, revenues can be either higher or lower in first and second price auctions. Moreover, asymmetric first-price auctions involve a tendency for weak bidders to bid more aggressively. Thus, although an equilibrium exists, the outcome may no longer be efficient, since the good is not always assigned to the bidder with the highest valuation.

its implications for the strategic scope an auctioneer may have once the auction format has been set. In a setting, where she has some means at hand to manipulate the valuations of the bidders after she committed to an auction mechanism, we analyze how she optimally acts to maximize her expected revenue.

Since the seminal contributions in auction theory, e.g. Vickrey (1961), Myerson (1981), or Riley and Samuleson (1981), an extensive strand of literature has emerged in this field of research. In addition to numerous papers dealing with the optimal design of auction mechanisms, much of the theoretical literature exhibits a strong focus on the behavior of bidders. Apart from optimal bidding strategies, topics range from bidder collusion to situations like discussed in Schwarz and Sonin (2005), where bidders can take investments or other private actions that affect their valuation.[2] In sharp contrast, the seller most commonly is assigned a merely passive role once the auction mechanism has been decided.[3] The main focus of this paper is to analyze what happens, if we allow the seller to take actions after the bidders signed up to the auction but before the bidding stage has started. More precisely, throughout the paper we maintain the assumption that the seller has full commitment with respect to the chosen allocation mechanism but can influence the bidders' valuations through another dimension of strategic choice.[4] Intuitively, by taking actions to support the weaker of two participating bidders, the seller can make them more competing rivals which in turn leads to an increase of her expected revenue. In this respect, it may then be optimal for the seller to favor a weaker bidder over the one who is ex ante most likely to win the auction. However, this results in ex post inefficiency in terms of social surplus, whenever strong bidder still wins and the seller "invested" in the ex post losing bidder.

The idea that supporting specific bidders may be beneficial to the auctioneer has been documented before. In an earlier study, Rothkopf et al. (2003) theoretically analyze public procurement auctions and find that subsidies to a class of relatively weaker competitors can lower the expected project cost to the government, since other bidders rationally respond by bidding more aggressively. Corns and Schotter (1999) conduct an experimental study for the case of affirmative action, where minorities are supported by the government. Against

[2]A comprehensive overview of the literature on auction theory is provided by e.g. Klemperer (2000).

[3]One exception is analyzed in Eso and Szentes (2007). In their model, the auctioneer has some superior information on a common-value good and can enter a signaling stage before the auction to share this information with the buyers.

[4]For an analysis of mechanism design problems with imperfect commitment see Bester and Strausz (2000).

the common argument that interference into the competitive process which prevents the most capable of being chosen must be wasteful and costly, they both argue that affirmative action may decrease the cost of government procurement under some circumstances, that is if the competition among bidders is sufficiently strong. McAfee and McMillan (1987) argue theoretically that preferential price treatment for designated bidders may be optimal in terms of revenue for the seller. Goeree and Offerman (2004) show that in some auction formats it can even be optimal for the auctioneer to award the highest losing bidder an ex-post premium that depends on the size of the bid she submitted. However, both subsidies and affirmative action programs are typically controversial and hard to justify, be it due to the need to raise (distortive) taxes to finance them or for political reasons. Moreover, all the above measures aiding designated bidders involve an direct interference with bids and the allocation mechanism and result in a mere redistribution of value. Our paper contributes to this strand of literature by demonstrating that auctioneers may have also other, more general means at hand to affect the bidders valuations. For example, prior to the actual auction stage a seller may add or alter some minor features of the item on sale, which asymmetrically affect the valuations of the bidders. In addition, we show that a seller can thereby exploit the beneficial effects of reducing asymmetries among bidders, but that this may cause inefficiencies with respect to social surplus. Say, for example, that prior to the execution of the auction the seller can add one or more features to the item on sale, which have no or very little intrinsic value themselves. However, they are assumed to asymmetrically affect the valuations of the bidders when added to the item, i.e. the bidders attach different values to these extra features.

To facilitate the analysis and to keep things tractable, we impose a number of simplifying assumptions on our model. As in much of the existing literature on asymmetric auctions, we assume there are exactly two bidders, one "strong" and the other "weak".[5] Hence, bidders are ex ante heterogeneous in the sense that their valuations are drawn from different distributions, such that one bidder is more likely to have a higher valuation than the other. Moreover, we consider second price auctions, where even with asymmetries present bidding one's true valuation is a dominant strategy for each bidder. Finally, we assume that the item is sold with certainty to one of the two bidders in the sense that the seller cannot

[5]The assumption of only two bidders is a convenient way to simplify notation and proofs. Moreover, the situation in this setting could also be interpreted similar to the second stage in Klemperer's (1998) *Anglo Dutch Auction*, where the two bidders are those who remain after a first round of a two stage auction process consisting of a final sealed-bid stage included into an otherwise-ascending auction.

credibly commit to not selling the item. Hence, a reserve price is non-credible in this setting in the sense that the bidders will anticipate that some kind of renegotiation or auction-like mechanism will subsequently be used to allocate the item.[6] These assumptions permit us to analyze the optimal strategy for the seller, when she can influence the bidders' valuations, and thus the degree of asymmetry among them.

This paper proceeds as follows. The basic setting of the model is presented in Section 5.2 and, following the lines of Cantillon (2000), the general effect of bidder asymmetries on expected auction revenue relative to a symmetric benchmark is discussed. Section 5.3 analyzes a simple scenario, where the seller has some fixed endowment she can use to individually support each of the two asymmetric bidders. Her choice will affect the latters' valuations and thereby also has consequences for the expected revenue she gains from the auction. A possible real world application for such a situation is discussed, before we turn to a scenario including the issues of standard setting and compatibility in Section 5.4. There, the basic model is extended to the case where the sellers action is costly and moreover may impose a negative effect on one of the bidder's valuation. Section 5.5 concludes. An Appendix collects formal proofs and provides an extension of the model.

5.2 Model Setting and Effect of Asymmetries

We focus on the case of two potential buyers $i \in \{w, s\}$ bidding for one unit of an indivisible object. The value bidder i attaches to the object is represented by v_i, and we assume that the independent-private values paradigm applies. While the valuation is private information to the individual bidder, it is common knowledge that v_i is independently drawn from the continuously differentiable atom-less cumulative distribution function F_i with support on $[\underline{v}_i; \overline{v}_i]$. Thus, the seller is able to identify different types of bidders, though not the actual valuation of any particular bidder. Define $S_i = \overline{v}_i - \underline{v}_i$ as the *spread* of the support for bidder i. The corresponding probability density function (PDF), $f_i(v) = F_i'(v)$, is finite and bounded away from zero. Moreover, assume that both bidders are risk neutral.

[6]See also Kirkegaard and Overgaard (2005).

Assumption 1 (Information:) (i) *Each of the two bidders knows the rules of the auction that the seller has chosen (and committed herself to).* (ii) *Bidder i knows her own valuation v_i, $i \in \{w, s\}$.* (iii) *Bidders' risk attitudes and the probability distributions of their valuations are common knowledge.*

In general, the two bidders are regarded as heterogeneous or *asymmetric* if $F_w \neq F_s$. To put more structure on the asymmetry, we employ the assumption that the value distributions of the bidders can be ranked according to first order stochastic dominance, which is standard in the theoretical literature on asymmetric auctions. Thus, for any realization v, one bidder has a higher probability of receiving an outcome equal to or better than v than the other.[7] Without loss of generality, suppose that F_s first order stochastically dominates F_w, that is $F_w(v) \geq F_s(v)\ \forall\ v$. Hence, there exists a "strong" bidder (s) who has a comparative advantage over the "weak" bidder (w) with respect to their valuations and winning probabilities. Note that first order-stochastic dominance applies if we assume that the bidders draw their valuations from distributions with different but intersecting supports.[8]

Assumption 2 (Bidder w is weak, bidder s is strong:) *Suppose that the boundaries of the supports are such that $\underline{v}_w \leq \underline{v}_s < \overline{v}_w \leq \overline{v}_s$.*

By Assumption 2, we have that $F_w(v) \geq F_s(v)\ \ \forall\ \ v \in (\underline{v}_w; \overline{v}_s)$. The first auction mechanism analyzed is a second-price sealed-bid auction, which is highly appealing due to its analytical simplicity. Moreover, since the highest-valuation bidder will be the one who actually wins the auction, the outcome is still efficient when bidders are asymmetric.[9] This allows us to highlight the phenomena caused by the existence of a strategic seller action during the course of the auction without getting too entangled in complex bidding mechanics.

Lemma 1 *In a second-price sealed bid auction, it is a weakly dominant strategy to bid one's own valuation. These bidding strategies are unaltered by the introduction of asymmetries.*

Proof. See Appendix 5.6.1.1. ■

[7]Note that under the given assumptions on the boundaries of the supports and the underlying distributions, also the stronger assumption of reverse hazard rate dominance ($\frac{h_s}{1-H_s} \leq \frac{h_w}{1-H_w}\ \forall\ v$) is satisfied, which is also often imposed and implies first-order stochastic dominance.

[8]See Figure 5.2a in Section 5.3. Also note that Assumption 2 ensures that supports are non-nested. In that case, hazard rate dominance and hence first-order stochastic dominance would be violated.

[9]Efficiency is not necessarily maintained in an asymmetric first price auction, since the object is assigned to a bidder other than the one with the highest valuation with positive probability. See e.g. Krishna (2002).

An immediate implication of Lemma 1 is that distribution of the selling price in a sealed-bid second price auction is given by:[10]

$$D(P) = Pr[min\{v_w, v_s\} \leq P] = F_w(P) + F_s(P) - F_w(P) \cdot F_s(P)$$

where $P \in \{\underline{v}_w, \overline{v}_w\}$, because from Assumption 2 the maximum price that a bidder potentially has to pay from an ex ante perspective is equal to the maximum realization the weak bidder's valuation can take. Since it is optimal for the bidders to bid their true value in a second price auction, the winner is always the one with the highest valuation.Note that $F_w(P) \cdot F_s(P)$ is the cumulative distribution function (CDF) of the expected highest realization from (F_w, F_s). Moreover, for any realized valuation v_i of bidder i, bidder j's value will be higher with probability $(1 - F_j(v_i))$ for $i \neq j$. Hence, the price that accrues to the seller equals the expectation of the second highest value, which we denote by $v_{(2)}^{w,s}$:

$$ER(F_w, F_s) \equiv v_{(2)}^{w,s} = \int_{\underline{v}_w}^{\overline{v}_w} v(1 - F_s(v))f_w(v)dv + \int_{\underline{v}_s}^{\overline{v}_s} v(1 - F_w(v))f_s(v)dv, \tag{5.1}$$

where the first term on the RHS is the expected valuation of bidder w conditional on bidder s winning the auction, and vice versa for the second integral. Since the supports of the distributions for the two bidders overlap, we can replace the integral borders by the respective minimum and maximum possible realizations from (F_w, F_s). By Assumption 2, these are given by $\underline{v}_w$ and $\overline{v}_s$, respectively. Thus, we can rearrange equation (5.1) to obtain a more intuitive expression for the expected seller revenues:

$$\begin{aligned}
v_{(2)}^{w,s} &\overset{Ass.2}{=} \int_{\underline{v}_w}^{\overline{v}_s} v(1 - F_s(v))f_w(v)dv + \int_{\underline{v}_w}^{\overline{v}_s} v(1 - F_w(v))f_s(v)dv \\
&= \int_{\underline{v}_w}^{\overline{v}_s} vf_w(v)dv + \int_{\underline{v}_w}^{\overline{v}_s} vf_s(v)dv - \int_{\underline{v}_w}^{\overline{v}_s} vF_s(v)f_w(v)dv - \int_{\underline{v}_w}^{\overline{v}_s} vF_w(v)f_s(v)dv \\
&= \int_{\underline{v}_w}^{\overline{v}_s} v\left[f_w(v) + f_s(v)\right]dv - \int_{\underline{v}_w}^{\overline{v}_s} vd\left[F_s(v)F_w(v)\right] \\
&= \int_{\underline{v}_w}^{\overline{v}_s} v\left[f_w(v) + f_s(v)\right]dv - v_{(1)}^{w,s}, \qquad (5.2)
\end{aligned}$$

where the last term, $v_{(1)}^{w,s}$, denotes the expected highest order statistic, i.e. the ex ante expected highest valuation among the two bidders.

[10]See e.g. Krishna (2002).

Before we turn to an analysis of a potential strategic scope for the seller to affect the auction outcome to her advantage, it proves useful to demonstrate the effect of asymmetries on the seller's revenue in general. Therefore, we proceed along the lines of Cantillon (2000), who constructs a symmetric benchmark auction environment to compare it to the expected revenue in the asymmetric case. We adopt the terminology from her paper, where she defines a combination of cumulative distribution functions $\{F_i, F_j\}$ as a *configuration*. The symmetric benchmark she designs is a new symmetric configuration of CDFs for the valuations of the two bidders labeled $\{F, F\}$, which has the following properties: First, by construction the expected highest order statistic, that is the expected highest realized valuation, is the same for both environments. Hence, the expected potential social surplus ("the size of the pie") is the same under both configurations. Second, she assumes that also the distribution of social surplus, i.e. the distribution of the highest order statistic, is the same for both configurations. In other words, if v is the highest realization from $\{F_w, F_s\}$ with probability p, then with probability p, v will also be the highest realization in the symmetric configuration.[11]

Definition 1 *Given two cumulative distribution functions $F_w(v)$ and $F_s(v)$ with supports on $[\underline{v}_w, \overline{v}_w]$ and $[\underline{v}_s, \overline{v}_s]$, respectively, their corresponding symmetric benchmark distribution is defined by $F(v) = \sqrt{F_w(v) \cdot F_s(v)}$ $\forall$ v and has support on $[\underline{v}, \overline{v}]$, where $\underline{v} = max\{\underline{v}_w, \underline{v}_s\}$ and $\overline{v} = max\{\overline{v}_w, \overline{v}_s\}$.*

Definition 1 implies that the respective CDFs of the highest order-statistics are indeed the same for both configurations, i.e. $F(v)F(v) = F_w(v)F_s(v)$. Let $f(v) = F'(v)$ denote the probability density function of the symmetric valuation primitive. The expected revenue to the seller is thus given by

$$\begin{aligned} v_{(2)} &= \int_{\underline{v}}^{\overline{v}} v(1 - F(v))f(v)dv + \int_{\underline{v}}^{\overline{v}} v(1 - F(v))f(v)dv \\ &= 2\int_{\underline{v}}^{\overline{v}} v(1 - F(v))f(v)dv \\ &= 2\int_{\underline{v}}^{\overline{v}} vf(v)dv - 2\int_{\underline{v}}^{\overline{v}} vF(v)f(v)dv \\ &= 2\int_{\underline{v}}^{\overline{v}} vf(v)dv - v_{(1)}, \end{aligned} \tag{5.3}$$

[11] Definition 1 corresponds to Definition 1 in Cantillon (2000), p. 6. For a more extensive discussion on the properties and the construction of an appropriate benchmark refer to Cantillon (2008).

where, analogously to above, $v_{(1)}$ denotes the expected highest order statistic in the benchmark setting. Since by construction we have that $v_{(1)} = v_{(1)}^{w,s}$, by subtracting (5.2) from (5.3) the last term in both expressions cancels out. Moreover, since $\underline{v} = \underline{v}_w$ and $\overline{v} = \overline{v}_s$ the difference in expected revenues across the two configurations amounts to

$$\begin{aligned} v_{(2)} - v_{(2)}^{w,s} &= 2\int_{\underline{v}_w}^{\overline{v}_s} vf(v)dv - \int_{\underline{v}_w}^{\overline{v}_s} v\left[f_w(v) + f_s(v)\right]dv \\ &= 2\int_{\underline{v}_w}^{\overline{v}_s} vf(v)dv - \int_{\underline{v}_w}^{\overline{v}_s} v\left[f_w(v) + f_s(v)\right]dv \\ &\overset{IBP}{=} -2\int_{\underline{v}_w}^{\overline{v}_s} F(v)dv + \int_{\underline{v}_w}^{\overline{v}_s} \left[F_w(v) + F_s(v)\right]dv \\ &= \int_{\underline{v}_w}^{\overline{v}_s} \left[\sqrt{F_w(v)} + \sqrt{F_s(v)}\right]^2 dv > 0, \end{aligned} \tag{5.4}$$

where the last steps are obtained through integration by parts (IBP). Since the expected highest order statistic is the same for both configurations, i.e. the maximum social surplus that can be attained, the strictly positive difference implies that the share of social surplus that the seller captures is reduced by asymmetry.

Proposition 1 *The expected revenue from a second-price auction for any asymmetric configuration (F_w, F_s) is lower than that under its corresponding symmetric benchmark (F, F).*[12]

Having laid out the general result that asymmetries hurt the auctioneer in terms of expected revenue from a second-price auction, we now turn to the question that is at the core of our interest: What is the optimal strategy for the seller, if she has some instrument at hand that allows her to influence the bidders' valuations - and hence the degree of asymmetry - during the course of the auction? For this purpose it proves useful to rearrange the expression for the auctioneers expected revenue. By Assumption 2, and since we have that $F_s(\cdot) = 0 \quad \forall\, v < \underline{v}_s$ and $F_w(\cdot) = 1 \quad \forall\, v \geq \overline{v}_w$, equation (5.1) can be written as

$$v_{(2)}^{w,s} = \int_{\underline{v}_w}^{\underline{v}_s} vf_w(v)dv + \int_{\underline{v}_s}^{\overline{v}_w} v(1 - F_s(v))f_w(v)dv + \int_{\underline{v}_s}^{\overline{v}_w} v(1 - F_w(v))f_s(v)dv. \tag{5.5}$$

[12]Proposition 1 recapitulates Theorem 1 in Cantillon (2008). In her paper, she shows that this result holds under very general conditions for a wider class of mechanisms than second-price auctions and under relaxed assumptions on the construction of the symmetric benchmark.

The first term captures the expected revenue whenever the realization of bidder w lies below $\underline{v}_s$, such that bidder s will win the auction for sure. Conditional on the expected second highest value falling into the region where the supports of the two distribution functions overlap,[13] the latter two terms describe the expected price that bidder w (or s, respectively) has to pay if she has the highest valuation. Furthermore, in what follows we relax the assumption from Definition 1 of a constant expected highest order statistic and focus on the maximization of expected revenue from an absolute perspective rather than in terms of the share of social surplus captured by the seller.

5.3 Scenario I: A Simple Model of Seller Interference

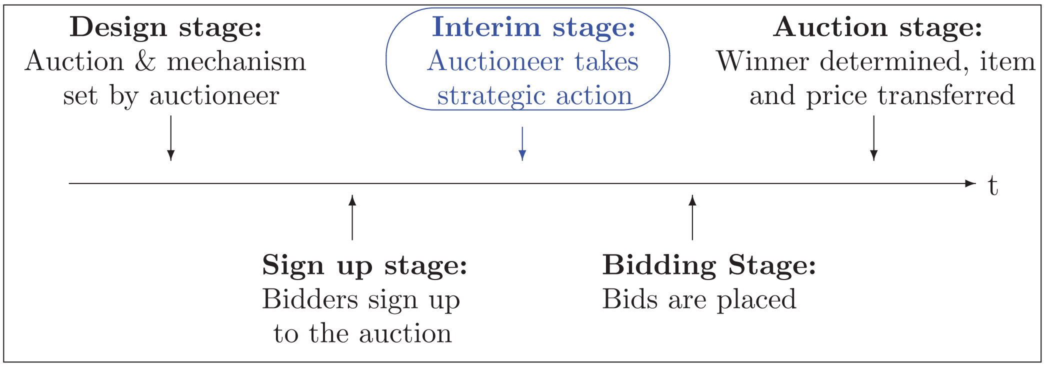
In most of the theoretical literature on auctions, the seller is assigned a passive role once the auction mechanism and specifications (e.g. the level of reserve price) have been set. In contrast to that, the main idea of this paper is that it is reasonable to assume that in some situations circumstances may arise that allow the seller to actively influence the bidders' valuation before the auction is actually carried out. Say, for example, that prior to the execution of the auction the seller can add or alter some features of the item on sale. While these may have no or very little intrinsic value themselves, they potentially affect the valuations of the bidders when added to the item, i.e. the latter possibly attach different values to these extra features. For instance, think of an expert and a layman bidding for a complex machine used to produce of some good, and that the latter has only very limited knowledge on how to operate the machine. In that case, an announcement of the seller during the course of the auction that he will provide technical assistance after the sale is likely to affect the laymen's valuation, while the expert is indifferent to this additional service, or "feature". In this section we will explicitly allow for such a possibility, where the seller has means to take impact on the distribution of valuations and ask for her optimal strategy.

[13] See the graph to the left of Figure 5.2a in the next section.

5.3.1 A Cake of Size x to Distribute

Within the standard framework presented in Section 5.2, the timing of events is as follows (see Figure 5.1): First, the auctioneer announces the auction for the item on sale, implements the design, and commits to the chosen mechanism. After this stage, the seller is assumed to remain passive. In the next step, the potential buyers - knowing their valuations - sign up for participation in the auction and place their bids. Finally, the auction is carried out, the winner is allocated the item and pays the price determined by the auction.

Figure 5.1: Timing of Events

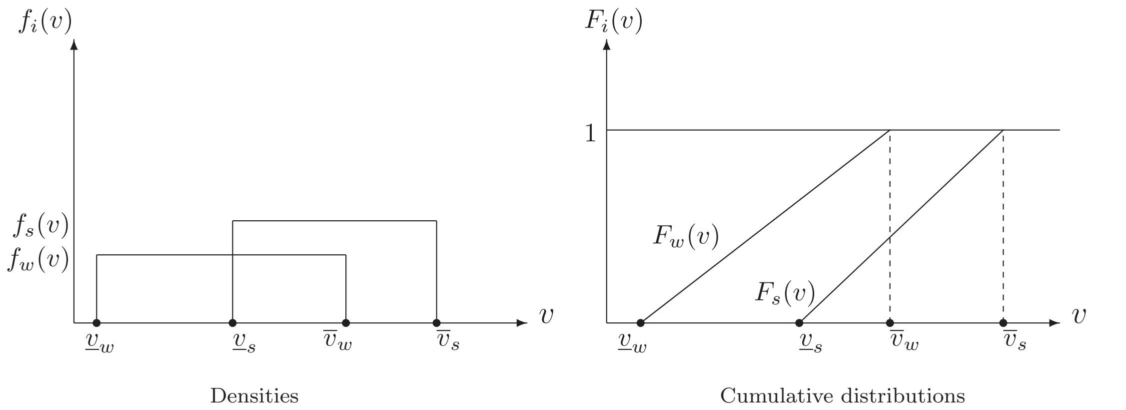


The novel feature of this paper, however, is to introduce an *interim stage* after the bidders have signed up, but before bidding takes place. During this stage, we assume that there exists scope for the seller to pursue a strategic action that affects the bidders' valuations.

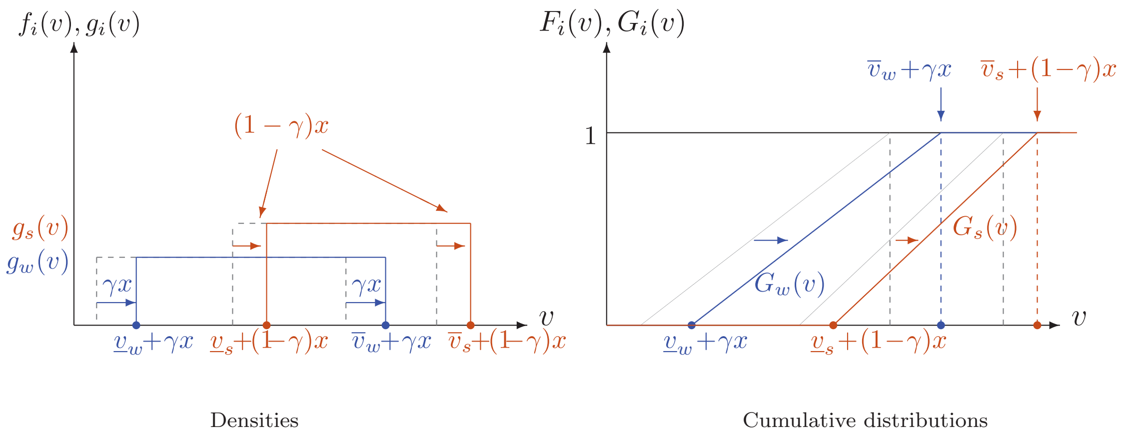
To start with, we concentrate on a stylized setting as a benchmark. Suppose that, in addition to the object on sale, the auctioneer has a divisible cake of size x which is valuable to the bidders but has no value to herself. In the interim stage, she can decide to split up this cake and allocate any fraction $\gamma \in [0,1]$ to bidder i, and $(1-\gamma)$ to bidder j. From the auctioneer's perspective, a non-negative allocation of a share of x to any of the bidders will result in an upward shift of the latters' value distribution: The weak bidders support shifts to the right by the additional value he gets, γx, and by $(1-\gamma)x$ for the strong bidder, respectively.[14] In general terms, the resulting CDFs and PDFs after the shifts will depend on the actual choice of γ and therefore we denote them by $G_i(v,\gamma)$ and $g_i(v,\gamma)$, respectively, for $i \in w, s$.

[14]Alternatively, instead of shifting the supports, one could also imagine a slightly different model, where the auctioneer's action puts stronger probability weight on the region of higher realizations while the original supports are maintained. Since this also increases the expected second highest value, and thus the expected revenue, we would get qualitatively similar results. However, for the sake of simplicity, we focus on shifts in the support.

Figure 5.2: Shift of Supports by Strategic Seller Action



(a) Initial densities and cumulative distributions



(b) Shift of supports by strategic seller action

Figure 5.2a depicts the original supports and Figure 5.2b illustrates the effect of the seller's choice for two asymmetric uniform distributions. Note that both the expected highest order statistic and the degree of asymmetry are affected by these shifts: While in either case the expected highest valuation increases, the maximum possible social benefit, i.e. the highest possible valuation, is only increased if the strong bidder receives a non-negative share ($\gamma < 1$). At the same time, the seller's expected revenue is equivalent to the expectation of the second highest value. Therefore, depending on the size of x relative to the degree of asymmetry among the bidders, from her perspective it might be beneficial to fully allocate x or at least a large fraction of it to the weak bidder. Intuitively, a reduction in the degree of asymmetry leads to more intense competition among the bidders. Importantly, after the seller has chosen the allocation, the situation is *as if* the valuations were drawn from the shifted distributions, even if the true initial valuations of the bidders are already realized.

Note further that bidding behavior is unaltered by the events in the interim stage: After the action took place it is still a (weakly) dominant strategy for each bidder to bid her true valuation. Hence, for any given cake size x, in the interim stage the auctioneer chooses the share to the weak bidder γ as to maximize the expected revenue from the auction, i.e.

$$\max_{\gamma}\ v_{(2)}^{w,s} = \int_{\underline{v}_w+\gamma x}^{\underline{v}_s+(1-\gamma)x} v g_w(v) dv + \int_{\underline{v}_s+(1-\gamma)x}^{\overline{v}_w+\gamma x} v(1-G_s(v,\gamma))g_w(v,\gamma)dv + \int_{\underline{v}_s+(1-\gamma)x}^{\overline{v}_w+\gamma x} v(1-G_w(v,\gamma))g_s(v,\gamma)dv, \quad (5.6)$$

The first order condition of the above expression with respect to γ implicitly determines the optimal allocation:[15]

$$\begin{aligned}
\frac{\partial v_{(2)}^{w,s}}{\partial \gamma} = & \int_{\underline{v}_w+\gamma x}^{\underline{v}_s+(1-\gamma)x} v g_w'(v,\gamma)dv + \int_{\underline{v}_s+(1-\gamma)x}^{\overline{v}_w+\gamma x} v\left[(1-G_s(v,\gamma))\cdot g_w'(v,\gamma) - G_s'(v,\gamma)\cdot g_w(v,\gamma)\right]dv \\
& + \int_{\underline{v}_s+(1-\gamma)x}^{\overline{v}_w+\gamma x} v\left[(1-G_w(v,\gamma))\cdot g_s'(v,\gamma) - G_w'(v,\gamma)\cdot g_s(v,\gamma)\right]dv \\
& - g_w(\underline{v}_w+\gamma x,\gamma)\cdot(\underline{v}_w+\gamma x)\cdot x + g_w(\overline{v}_w+\gamma x,\gamma)\cdot(\overline{v}_w+\gamma x)\cdot(1-G_s(\overline{v}_w+\gamma x,\gamma))\cdot x \\
& + g_s(\underline{v}_s+(1-\gamma)x,\gamma)\cdot(\underline{v}_s+(1-\gamma)x)\cdot(1-G_w(\underline{v}_s+(1-\gamma)x,\gamma))\cdot x = 0. \quad (5.7)
\end{aligned}$$

On this general level, we can only infer that the optimal allocation choice depends on the properties of the underlying distribution functions from which the bidders' valuations are realized and their reaction to changes in γ. To be able to derive more intuitive insights, we additionally impose the simplifying assumption that the valuation primitives follow uniform distributions of the form

$$G_w(v,\gamma) = \begin{cases} 0 & \text{if } v \leq \underline{v}_w + \gamma x \\ \frac{v-(\underline{v}_w+\gamma x)}{\overline{v}_w-\underline{v}_w} & \text{if } \underline{v}_w + \gamma x < v < \overline{v}_w + \gamma x \\ 1 & \text{if } v \geq \overline{v}_w + \gamma x \end{cases} \qquad \text{and}$$

$$G_s(v,\gamma) = \begin{cases} 0 & \text{if } v \leq \underline{v}_s + (1-\gamma)x \\ \frac{v-(\underline{v}_s+(1-\gamma)x)}{\overline{v}_s-\underline{v}_s} & \text{if } \underline{v}_s + (1-\gamma)x < v < \overline{v}_s + (1-\gamma)x \\ 1 & \text{if } v \geq \overline{v}_s + (1-\gamma)x \end{cases}$$

for the weak and the strong bidder, respectively.[16] The corresponding densities are given by

[15] See Appendix 5.6.1.2 for the details of the calculation.

[16] We focus on uniform distributions for the sake of tractability. In Appendix 5.6.2 we present a numerical simulation employing doubly truncated normal distributions yielding qualitatively similar results.

$$g_w(v,\gamma) = \begin{cases} \frac{1}{\overline{v}_w - \underline{v}_w} & \text{if } \ \underline{v}_w + \gamma x < v < \overline{v}_w + \gamma x \\ 0 & \text{otherwise} \end{cases} \qquad \text{and}$$

$$g_s(v,\gamma) = \begin{cases} \frac{1}{\overline{v}_s - \underline{v}_s} & \text{if } \ \underline{v}_s + (1-\gamma)x < v < \overline{v}_s + (1-\gamma)x \\ 0 & \text{otherwise} \end{cases}$$

Conveniently, the functional form of the uniform densities is unaltered by the shifts in the support. Hence, we have that $\frac{\partial g_i(\cdot)}{\partial \gamma} = 0$. Moreover, this specification allows us to derive a closed form solution for the optimal allocation from the first order condition.[17] In particular, the optimal share to the weak bidder, γ, must satisfy

$$\gamma = \frac{1}{2} + \frac{-(\overline{v}_w - \underline{v}_s) + \sqrt{S_s \cdot S_w}}{2x}, \tag{5.8}$$

where $S_s = \overline{v}_s - \underline{v}_s$ and $S_w = \overline{v}_w - \underline{v}_w$. Intuitively, the term $-(\overline{v}_w - \underline{v}_s) + \sqrt{S_s \cdot S_w}$ can be interpreted as a measure for the *degree of asymmetry* among the two bidders. From this relation, we can derive a first set of results. First, as one would expect from the analysis in Section 5.2, whenever the bidders draw their valuations from the same distribution, an equal split of the cake x among the bidders is optimal. In other words, in the absence of asymmetry the seller will not favor any bidder over the other in the auction process.

Proposition 2 (Symmetric bidders) *For $\underline{v}_w = \underline{v}_s$ and $\overline{v}_w = \overline{v}_s$ the optimal allocation choice is given by $\gamma = \frac{1}{2}$.*

Proof. When bidders are symmetric, we have that $\underline{v}_s = \underline{v}_w \equiv \underline{v}$ and $\overline{v}_s = \overline{v}_w \equiv \overline{v}$. Hence, the numerator in the last term of (5.8) becomes $\underline{v} - \overline{v} + \sqrt{(\overline{v} - \underline{v})^2} = 0$, giving the result. ■

Second, as long as the shift in the support is not sufficient to outweigh the relative asymmetry among the bidders, it is always optimal to allocate the full amount of x to the weak bidder. In line with the intuition from Section 5.2 above, reducing the asymmetry has a positive effect on expected revenue. Conversely, if the cake size is large enough to fully compensate for the asymmetry, both bidders receive a positive share from x, with over-proportional weight on the weaker bidder ($\gamma > \frac{1}{2}$). Third, the optimal share to w ceteris paribus increases in the spreads of the supports (S_s and S_w), and decreases with the length of the intersection region

[17] See Appendix 5.6.1.3 for the detailed derivation.

$(\overline{v}_w - \underline{v}_s)$. Intuitively, the larger the comparative advantage of the strong bidder, the larger are the benefits to the seller from supporting the weaker one. Proposition 3 summarizes these findings.

Proposition 3 (Allocation choice) *(i) If $x \leq -(\overline{v}_w - \underline{v}_s) + \sqrt{S_s \cdot S_w}$ the optimal allocation is a corner solution at a share of $\gamma = 1$. (ii) For $x \geq -(\overline{v}_w - \underline{v}_s) + \sqrt{S_s \cdot S_w}$ the seller sets $\frac{1}{2} \leq \gamma < 1$, and bidders are symmetric ex-post. If x becomes large, the optimal allocation converges to an equal split in the limit. (iii) The share to the weak bidder increases in the spreads of the supports, $\frac{\partial \gamma}{\partial S_s} = \frac{\partial \gamma}{\partial S_w} = \frac{1}{4x\sqrt{S_w \cdot S_s}} > 0$, and decreases in the length of their intersection region, $\frac{\partial \gamma}{\partial(\overline{v}_w - \underline{v}_s)} = -\frac{1}{2x} < 0$.*

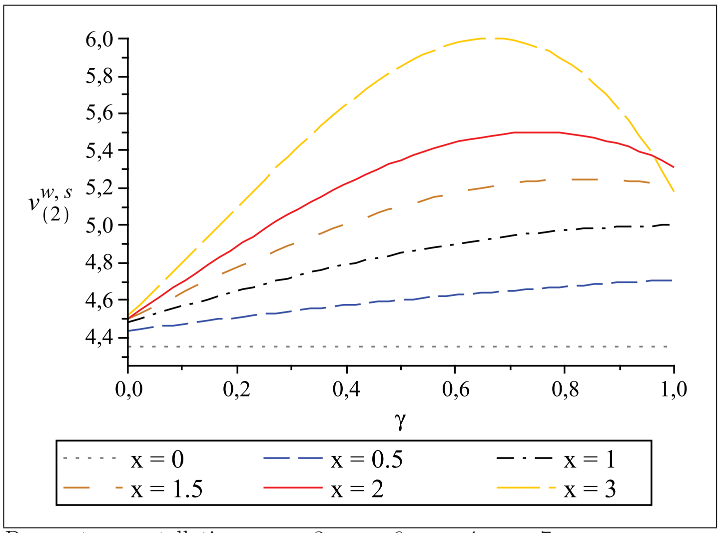
Proof. For result (i) it suffices to show that $\gamma = 1$ if the last term on the RHS of (5.8) is greater than or equal to $\frac{1}{2}$. This is the case if and only if the above condition holds, i.e. $x \leq -(\overline{v}_w - \underline{v}_s) + \sqrt{S_s \cdot S_w}$. The reverse argument applies to the first part of result (ii), respectively. The second part of (ii) is a direct implication of result (i): Whenever the degree of asymmetry is smaller than x, the auctioneer optimally allocates the cake such that symmetry among bidders is restored. All remaining results are directly implied by the partial derivatives of the optimality condition w.r.t. S_i for $i = \{w, s\}$ and $(\overline{v}_w - \underline{v}_s)$. ■

To emphasize these findings, Figure 5.3 depicts the results from a numerical simulation within the uniform specification.[18] For the given set of parameters, the expected revenue function in relation to the share γ for different cake sizes is depicted. Consistent to above, the figure shows that for x sufficiently small $v_{(2)}^{w,s}$ takes its maximum at $\gamma = 1$ (corner solution). If the impact of the seller action gets large relative to the degree of asymmetry, the optimal share to the weak bidder becomes smaller and also the strong bidder is allocated a non-negative share of the cake. In the limit, γ converges to $\frac{1}{2}$, confirming our above conjecture.

Summarizing, if the seller has an action to influence the bidders during the course of the auction, she will exclusively favor the weak bidder whenever the impact of her action is smaller than the degree of asymmetry among the bidders. By doing so, she can make them more competing rivals, leading to an increase in the expected second highest value, and hence her expected profit.

[18] The propositions have proven to be valid for any parameter constellation satisfying Assumption 2 and even carry over to some cases where the supports of the bidders are nested (though the first-order stochastic dominance property no longer applies). Moreover, they hold for a wide class of distribution functions. See Appendix 5.6.2 for an analogous simulation involving truncated normal distributions.

Figure 5.3: Optimal Allocation (Uniform)



Parameter constellation: $\underline{v}_w = 3, \overline{v}_w = 6, \underline{v}_s = 4, \overline{v}_s = 7$.
Relative degree of asymmetry: $-(\overline{v}_w - \underline{v}_s) + \sqrt{S_s \cdot S_w} = 1$.

To simplify the analysis, up to now we imposed the assumption that the seller's decision is merely an allocation choice and free of cost. Thus, in this simple setting the allocation choice is unambiguously revenue enhancing, since otherwise the seller could always opt to forego her action and dispose x.[19] In Section 5.4 we relax this restriction and turn to the analysis of a model, where we introduce costly actions. Before that, however, we briefly discuss a real world application, where a scenario similar to that presented in this section might arise.

5.3.2 A Real World Application

One potential application for the above scenario can be found in the context of the sale of state-owned energy-sector assets. More precisely, we consider privatizations of the gas-distribution network. In its Guideline 2003/55/EG the EU-Commission advocates the liberalization of network-industries as a way to increase efficiency, competition, and service quality at reduced prices. Even earlier, in 2000, it called for a fast realization of a single European energy market and devised a schedule for obligatory stepwise liberalization. As a consequence, a large number of privatizations was observable in the recent years within current and prospective EU-member countries.[20]

[19]Though this outside option is not explicitly included in the model, is straightforward to show that expected revenue is increasing in x.

[20]A detailed overview of the European policies and guidelines regarding the gas and energy markets is available at http://ec.europa.eu/energy/gas_electricity/index_en.htm.

This required liberalization clearly involves several objectives for national governments. First, one target of privatization is to yield a high price for tendering the assets at stake. Second, to maintain the security of energy supplies, national governments might favor the creation of a national-champion firm, which is able to take a strong strategic position in the future international competition with other global players and has its roots within the country.[21] While the first goal can be achieved by a well-designed selling mechanism like an auction, the second objective is more complex since the EU imposes a *non-discrimination principle* for public tenders. Hence, neither subsidies to specific bidder firms nor affirmative actions are eligible in the tender process.

However, in this industry large economies of scale can emerge, for instance from increasing total uninterrupted pipeline length or extending the capacity diameter of the pipes.[22] Thus, as the imposed liberalization plan involves a stepwise privatization, the choice of the selling sequence may create an alternative instrument for the national governments. The intuition for this is that a network distributor may derive additional benefits from scale economies, whenever she manages to purchase a sector that has a *direct connection* to the pipeline network it already owns.[23]

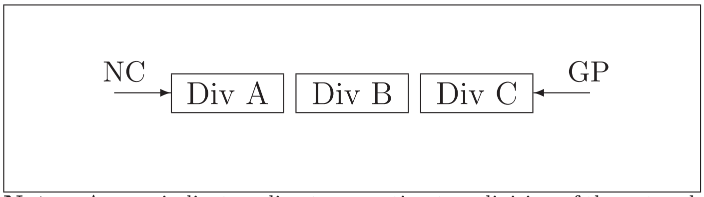
Translated to our model setting, consider a small national firm (NC) and a strong global player (GP), where both are interested in acquiring a part of a country's gas-distribution network. It is reasonable to assume that NC is financially weaker than GP, i.e. draws its valuation from a distribution with lower support in our terminology. Suppose that the gas distribution network in state ownership consists of three separate divisions, where the geographic placement is such that NC (GP) owns parts of the network in the direct neighborhood of division A (C), as illustrated in Figure 5.4. For the remaining division no such direct connection exists. Thus, in this configuration the "cake" that can be allocated by the seller (the government) can be thought of the synergy-effects that potentially can accrue to one of the bidders, when a specific sector is put on sale first.

[21] One recent example for such intentions of national governments can be seen in the case of Endesa. After the German company E.ON made an offer to takeover the Spanish energy supplier, the Spanish government intervened, stating it wanted Endesa to remain under "Spanish ownership". Even though the EU-commission instituted proceedings against the governmental intervention, finally the debate ended with E.ON redrawing its offer after fruitless negotiations.

[22] For a technical approach to the costs of gas distribution networks refer to e.g. Yepez (2008). An empirical study to estimate the components driving the cost of pipeline operation is provided in Bernard et al. (2002).

[23] Note that in line with our assumptions the government derives no direct value from the synergy effects.

Figure 5.4: Network Structure and Direct Connections



Notes: Arrows indicate a direct connection to a division of the network in state ownership. NC (GP) has a direct connection to division A (C), while no direct connection exists for division B.

Suppose the government has committed to a sealed-bid second-price-auction. As discussed above, it cannot take any direct measures to favor the NC. However, by selling division A first it can implicitly allocate additional benefits from increased economies of scale to the national firm in case it should win the auction, which will not accrue to GP. As a consequence, the relative degree of asymmetry among the two competitors is reduced.[24] This way, the government achieves both increased auction revenues and a better position for the NC, even if it may be still the case that GP is most likely to win.

We perceive this scenario as a nice illustration of the main motivation behind this paper, namely that there may be very indirect though important channels through which a seller can take influence on the bidders valuations, which should not be neglected. However, given the complex nature of network industries and the energy sectors, this example is clearly an over-simplification of the reality and neglects numerous crucial aspects. To name only a few examples, tender offers typically involve other standards than second-price auctions, vertical integration of service providers and network operators makes (legal) ownership-unbundling an important topic, and pre-auction offers to the candidates most likely to win are not unusual despite the restrictive EU-regulation.

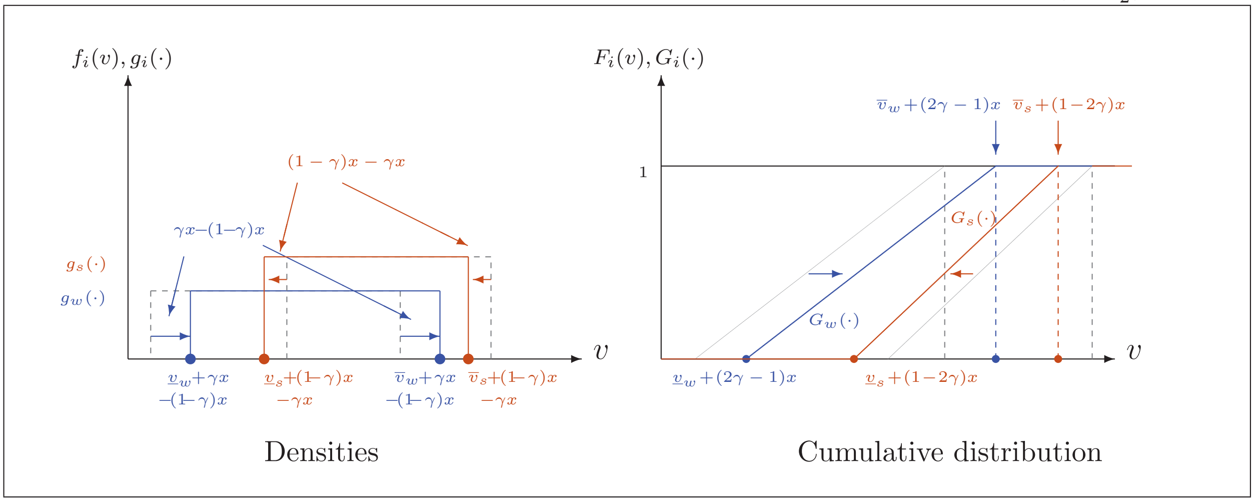
5.4 Scenario II: Bidders with Diverging Tastes

Until now we only considered costless actions with an unambiguously positive effect on the bidders' valuations. In this section, we now turn to the case where the seller's choice is costly. These costs are assumed to be twofold: First, the action involves a direct implementation cost for the seller, which accrues upon its execution. Second, we introduce a second dimension of

[24]Note that this scenario will produce a corner solution, since a firm gets either benefits from the synergy or not. For simplicity, we assume that the value from the synergy effect is the same across bidders.

asymmetry among the two bidders. Suppose they are not only asymmetric with respect to their distributions of value, but also differ in their taste for a certain specification of some feature or *standard* of the item on sale. In other words, favoring bidder i may trigger a negative impact on the valuation of bidder j. This is illustrated in Figure 5.5 for the case of uniform distributions.

Figure 5.5: Shift of Supports if Seller Moves Toward Standard of Weak Bidder ($\gamma > \frac{1}{2}$)



Analogously to above, assume that a seller faces a "weak" and a "strong" bidder interested in an item sold via a second-price auction. Before the bidding stage, the seller can decide to implement a change of the standard γ at cost $C(\gamma)$, where $C(\cdot)$ is convex and γ is continuous on the interval $[0, 1]$. Without loss of generality, assume that standard $\gamma = 0$ is preferred by the strong bidder, while bidder w has a preference for standard $\gamma = 1$. If no investment is taken, the initial standard of $\gamma = \frac{1}{2}$ is kept, and no costs accrue to the seller with $C(\frac{1}{2}) = 0$ and $C'(\frac{1}{2}) = 0$. Suppose the seller implements a standard closer to the one preferred by bidder w, that is $\gamma > \frac{1}{2}$. In that case, we assume that the latter gains an additional utility of $(2\gamma - 1)x > 0$ from increased complementarity, while bidder s experiences a disutility of $(1 - 2\gamma)x < 0$. Vice versa, if the seller instead moves toward the other extreme, the inequalities are reversed.[25]

For example, the standards in our model could reflect rival software applications for enterprise resource planning. Suppose a firm is target for takeover by two asymmetric competitors, w

[25]For the ease of exposition, we treat the maximum complementarity benefits x equally for both bidders. If this assumption is relaxed, it is straightforward to show that the tendency of the seller to favor the weak bidder becomes stronger whenever $x_w > x_s$. For $x_w < x_s$ there exists a cut-off value at $\frac{x_s}{x_w} \equiv z$, below which the auctioneer still moves towards w's preferred standard. The level of z depends on the parameter specification of the valuation primitives, the supports, and the cost function.

and s. Say, the weak bidder uses SAP-software for its daily business, while bidder s works with Oracle, which exhibit only a limited degree of compatibility to each other.[26] Assume further, that the target firm initially operates some third software, that exhibits an equal degree of compatibility to both. However, incurring some implementation cost, the firm can switch parts of its system towards the preferred standard of one of the bidders before the actual auction stage starts. More general, one could think of operating systems, technical standards, organizational topics, and many other features, for which the bidders might differ in their preferences.

Analytically, the seller chooses the standard as to maximize his expected revenues net of the implementation costs, $\Pi = v_{(2)}^{w,s} - C(\gamma)$:

$$\max_{\gamma} \int\limits_{\underline{v}_w+(2\gamma-1)x}^{\underline{v}_s+(1-2\gamma)x} v g_w(v) dv + \int\limits_{\underline{v}_s+(1-2\gamma)x}^{\overline{v}_w+(2\gamma-1)x} v[(1-G_s(v,\gamma))g_w(v,\gamma)dv + (1-G_w(v,\gamma))g_s(v,\gamma)]dv - C(\gamma), \tag{5.9}$$

where analogously to above $G_i(v,\gamma)$ and $g_i(v,\gamma)$ denote the CDFs and PDFs that result from the chosen standard γ (see Figure 5.5). Taking the derivative of (5.9) with respect to γ and rearranging yields the optimality condition.[27]

$$\begin{aligned}
\frac{\partial v_{(2)}^{w,s}}{\partial \gamma} = \quad & \int\limits_{\underline{v}_w+(2\gamma-1)x}^{\underline{v}_s+(1-2\gamma)x} v g_w'(v,\gamma) dv + \int\limits_{\underline{v}_s+(1-2\gamma)x}^{\overline{v}_w+(2\gamma-1)x} v\left[(1-G_s(v,\gamma))g_w'(v,\gamma) - G_s'(v,\gamma)g_w(v,\gamma)\right] dv \\
+ \quad & \int\limits_{\underline{v}_s+(1-2\gamma)x}^{\overline{v}_w+(2\gamma-1)x} v\left[(1-G_w(v,\gamma))g_s'(v,\gamma) - G_w'(v,\gamma)g_s(v,\gamma)\right] dv \\
- \quad & 2x \cdot g_w(\underline{v}_w + (2\gamma-1)x, \gamma) \cdot (\underline{v}_w + (2\gamma-1)x) \\
+ \quad & 2x \cdot g_w(\overline{v}_w + (2\gamma-1)x, \gamma) \cdot (\overline{v}_w + (2\gamma-1)x) \cdot (1 - G_s(\overline{v}_w + (2\gamma-1)x, \gamma)) \\
+ \quad & 2x \cdot g_s(\underline{v}_s + (1-2\gamma)x, \gamma) \cdot (\underline{v}_s + (1-2\gamma)x) \cdot (1 - G_w(\underline{v}_s + (1-2\gamma)x, \gamma)) \\
= \quad & C'(\gamma)
\end{aligned} \tag{5.10}$$

Intuitively, in the optimum the marginal benefit from moving towards one of the bidders' preferred standard equals the sum of marginal costs due to the downward shift of the other bidders distribution and the marginal implementation costs. To simplify the analysis, we

[26] In fact, until quite recently, Oracle followed a strategy of foreclosure to bind its customers and increase the obstacles to switch to rival enterprise resource planning vendors.

[27] Calculations are analogous to those presented in Appendix 5.6.1.2.

again focus on uniformly distributed valuation primitives. The CDFs and PDFs for the uniform specification are stated below. Note that the expressions are similar to the previous example except for the different effects on the boundaries of the supports.

$$G_w(v,\gamma) = \begin{cases} 0 & \text{if } \; v \leq \underline{v}_w + (2\gamma-1)x \\ \frac{v-(\underline{v}_w+(2\gamma-1)x)}{\overline{v}_w-\underline{v}_w} & \text{if } \; \underline{v}_w + (2\gamma-1)x < v < \overline{v}_w + (2\gamma-1)x \\ 1 & \text{if } \; v \geq \overline{v}_w + (2\gamma-1)x \end{cases}$$

$$G_s(v,\gamma) = \begin{cases} 0 & \text{if } \; v \leq \underline{v}_s + (1-2\gamma)x \\ \frac{v-(\underline{v}_s+(1-2\gamma)x)}{\overline{v}_s-\underline{v}_s} & \text{if } \; \underline{v}_s + (1-2\gamma)x < v < \overline{v}_s + (1-2\gamma)x \\ 1 & \text{if } \; v \geq \overline{v}_s + (1-2\gamma)x \end{cases}$$

$$g_w(v,\gamma) = \begin{cases} \frac{1}{\overline{v}_w-\underline{v}_w} & \text{if } \; \underline{v}_w + \gamma x < v < \overline{v}_w + \gamma x \\ 0 & \text{otherwise} \end{cases}$$

$$g_s(v,\gamma) = \begin{cases} \frac{1}{\overline{v}_s-\underline{v}_s} & \text{if } \; \underline{v}_s + (1-\gamma)x < v < \overline{v}_s + (1-\gamma)x \\ 0 & \text{otherwise} \end{cases}$$

Suppose further that the implementation costs take a quadratic form: $C(\gamma) = \lambda(\frac{1}{2} - \gamma)^2$, where $\lambda > 0$ is an exogenous cost parameter.[28] In this case a closed form solution to the auctioneer's problem can be derived. More precisely, the optimal standard is given by[29]

$$\gamma = \frac{1}{2} + \frac{8(\underline{v}_s - \overline{v}_w)x^2 - \lambda S_w S_s + \sqrt{S_w S_s \left(\lambda^2 S_w S_s + 16\lambda x^2(\overline{v}_w - \underline{v}_s) + 64x^4\right)}}{32x^3}, \quad (5.11)$$

where $S_s = \overline{v}_s - \underline{v}_s$ and $S_w = \overline{v}_w - \underline{v}_w$. This conditions allows us to infer several insights, which are summarized in Proposition 4.

Proposition 4 (Optimal standard under costly action) *(i) If bidders are symmetric, the optimal standard is given by $\gamma = \frac{1}{2}$. (ii) For x sufficiently small, it is optimal for the auctioneer to implement the standard preferred of the weak bidder at $\gamma = 1$. (iii) If the gains*

[28]Note that indeed $C(\cdot) = 0$ and $C'(\cdot) = 0$ if the original standard of $\gamma = \frac{1}{2}$ is maintained. Qualitatively, the results do not depend on the specific form of the cost function.

[29]The second-order condition is non-positive and thus the above condition describes indeed a local maximum. For the sake of brevity we omit the details of the purely mechanical but somewhat tedious derivation.

from complementarity ($x \to \infty$), or the implementation costs ($\lambda \to \infty$) become large, the optimal standard ceteris paribus converges to $\gamma = \frac{1}{2}$.

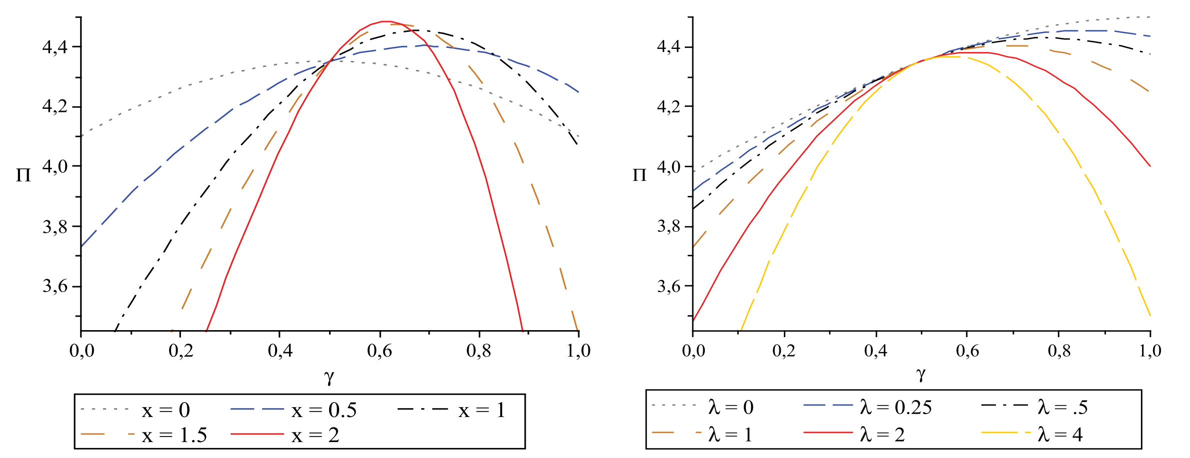
Proof. See Appendix 5.6.1.4. ■

Note further, that in absence of direct implementation costs ($\lambda = 0$), the first order condition reduces to

$$\gamma = \frac{1}{2} + \frac{\underline{v}_s - \overline{v}_w + \sqrt{S_w \cdot S_s}}{4x}, \tag{5.12}$$

which closely resembles the optimality condition (5.8) that was derived in Section 5.3. Hence, our previous results from Proposition 3 qualitatively carry over to this scenario. In particular, the second term on the RHS of (5.12) is smaller than the corresponding term on the RHS of (5.8) by factor $\frac{1}{2}$. Intuitively, while favoritism towards the weak bidder is still an optimal strategy, for any x the seller favor weak bidder relatively less compared to the case without the externality on bidder s.

Figure 5.6: Optimal Standard (Uniform)



(a) Revenue in relation to importance of standard (b) Revenue in relation to implementation cost

Parameter constellation: $\underline{v}_w = 3$, $\overline{v}_w = 6$, $\underline{v}_s = 4$, $\overline{v}_s = 7$, $\lambda = 1$ (Figure a), and $x = 0.5$ (Figure b).

Also for this setting we conduct numerical simulations to illustrate our findings. Holding the cost parameter λ constant, Figure 5.6a depicts the relation between the choice of standard γ and the expected profit for different values of complementarity gains, measured as the impact of x on the bidders' utility. We find that with increasing importance of the complementarity, the auctioneer becomes more reluctant to move away from its original value. In other words, if x becomes sufficiently large, even a small change of standard suffices to mitigate the effects of asymmetry. Conversely, as long as x is sufficiently small it is revenue enhancing to implement the weak bidder's preferred standard at $\gamma = 1$.

The simulation underlying Figure 5.6b analyzes how the auctioneer's optimal decision varies if the cost parameter λ increases for a given x. Clearly, if implementation costs are small compared to the gains from reduced asymmetry, it is optimal for the auctioneer to move towards bidder w's standard. For larger costs, the optimal choice converges to taking no action at all and sustaining the initial standard.

Summarizing, in terms of expected auction revenue it can be optimal for the auctioneer to move the standard towards the preference of the weak bidder. However, unlike the previous scenario, by doing so she hurts bidder s in this case, because the latter's maximal possible valuation is reduced if a standard less preferred to her is implemented. Moreover, if we interpret the strong bidders ex ante maximum valuation as the highest social surplus attainable, from a welfare perspective the seller's action causes an inefficiency as potential value is destroyed: $\overline{v}_s + (1 - 2\gamma) < \overline{v}_s$ for $\gamma > \frac{1}{2}$. Intuitively, in case the strong bidder wins the auction, his valuation is not only reduced indirectly by more competition on behalf of the weak bidder, but also directly through the externality caused by the seller's action. Thus, total social surplus is lower than without the auctioneer's interference.

5.5 Conclusion

We examine theoretically a setting where the auctioneer faces two asymmetric bidders. The novel feature is that we introduce a possibility that she can "strategically" influence the degree of asymmetry, and thus competition, by shifting the distribution of the latters' valuations. Our model provides two main insights. First, it is revenue enhancing for the seller to support the bidder who is originally the most likely to lose. The origin of this effect stems from the structure of a second-price auction: Since the revenues to the seller equal the second highest bid, the only way to improve upon her profit is to increase the expected second highest order statistic, i.e. the expected payment by the winning bidder. Furthermore, we show that this result holds true even if this favoritism causes a negative impact on the competing "strong" bidder. Second, we argue that this support may take place during an interim stage, when the auction mechanism is already set up and committed to, without violating the rules of the implemented auction. At a policy level, these results suggest that in merger analysis and public tenders involving auctions, interference with the competitive outcomes may occur through more subtle and indirect channels other than handicaps or affirmative action.

Our findings complement those of Bulow et al. (1999). They analyze the effect of one bidder owning a toehold in a company subject for takeover in a setting with (ascending) first-price auctions and common-values. Among other things, they show that the "*board of a target company may [...] wish to 'level the playing field' by selling a toehold to a new bidder [...]*" (Bulow et al., 1999, p.450). In other words, under some circumstances it can be beneficial in terms of revenue for the non-bidding shareholders to offer a toehold in form of a cheap (or even free) stake in the firm to another competing bidder. The intuition for this result is that "*a bidder that owns a toehold has an incentive to bid aggressively since every price it quotes represents not just a bid for the remaining shares but also an ask for its own holdings*" (Bulow et al., 1999, p.428). Hence, by allocating a toehold to a second bidder, the competitiveness among bidders – and thereby expected revenues from the auction – can be increased. Though their setting differs in several aspects, this finding is similar to ours if we interpret the "action" of the seller as the offer of a (cheap) stake in the item to the "weak" bidder, partially making him a *residual claimant* of the auction outcome.[30]

Another interpretation arises from a governance perspective. Intuitively, in our takeover example, the board's (i.e. the "auctioneer" in our terminology) action to favor the "losing bidder" may be falsely perceived as a poison pill by the shareholders, since the most likely candidate to win the auction is made worse off. However, as our model shows, it actually acts in their interest by maximizing expected revenue. Consider for example the case of the enterprise software provider PeopleSoft. During 2003 a possible takeover in the future was anticipated, presumably by its rival, Oracle Corporation. In reaction, PeopleSoft guaranteed a refund up to five times the cost of a product license to its customers in case the customer support was reduced within the next four years. Commonly, this measure was considered as a strategy to deter the takeover, because this would impede the expected plans of Oracle to phase out former Peoplesoft products in case of a successful acquisition. However, in the light of our model another interpretation arises. Suppose the takeover took an auction

[30] However, note that such an offer targeted exclusively to a specific bidder potentially collides with established competition regulations on non-discrimination.

format and that there was a (weaker) candidate other than Oracle to compete for PeopleSoft, who plans to continue the latter's product line after the takeover. Such a company would not necessarily be harmed by the refund guarantee. Conversely, it might even benefit from increased long term relationship with existing customers. As a result, the refund might be a way to exploit the revenue enhancing effect of reduced asymmetry.

Several extensions to this research suggest themselves. First, it would be interesting to generalize the idea of the model to other auction formats like the first-price or the English auction.[31] Second, our approach includes a number of simplifying assumptions on the environment which could be relaxed. For example, a generalization for arbitrary distribution functions and an extension to the N bidder case might be interesting. Third, we do not consider the welfare implications for the buyers, which might be also worthwhile to pursue. Finally, we abstract from the issue of reserve prices by restricting our attention to "must-sell" auctions, where a reserve price is not applicable for reasons laid out above. However, it might also be interesting to compare the effects of the seller action on expected revenue to a setting where the seller sets a non-negative reserve price. This is left to future research.

[31] Though the findings in Bulow et al. (1999) suggest that similar effects might persist in that case, the driving forces behind them are likely to differ substantially involving a fully-fledged analysis of bidder behavior and bidding strategies.

5.6 Appendix

5.6.1 Proofs and calculations

5.6.1.1 Lemma 1

The proof for the first part of the lemma is a standard textbook result (cf. Krishna, 2002, p.15). If bidder i bids his true valuation, she will win the auction whenever $v_i > p_j$ and she will lose if $v_i < p_j$ where $p_j = \max_{j \neq i} b_j$ is the highest competing bid. It remains to show, that deviating by bidding less (more) than his true valuation, i.e. $z_i < v_i$ ($z_i > v_i$), is not profitable.

First, consider the cases where either $p_j < z_i < v_i$ or $z_i < v_i < p_j$. For both cases, the auction outcome and profits are unaltered if she bids z_i instead of her valuation. However, if $v_i > p_j > z_i$ then bidder i loses the auction, whereas she would have won and experienced a positive surplus if she had bid v_i. Thus, deviating by bidding less than ones own valuation can never be profitable. In the same lines, bidding more than one's valuation can never be optimal. Given that in equilibrium all other j bidders play according to the bidding strategy $\beta(v) = v$, this is also a (weakly) dominant strategy for bidder i.

Moreover, the same arguments also apply for the second part of Lemma 1. Since the individual rationale of this proof does not depend on the valuation primitives but on the actual realizations of the bidders, the introduction of asymmetries does not affect the bidding behavior and it is still a weakly dominant strategy for each bidder to bid her true valuation.

■

5.6.1.2 General optimality condition (Section 5.3)

To derive the FOC from (5.6) with respect to γ, we proceed stepwise by applying the Leibniz rule for parametric integrals to each of the three integral terms separately. The derivative of the first term yields the following expression.

$$\int_{\underline{v}_w+\gamma x}^{\underline{v}_s+(1-\gamma)x} v g_w'(v,\gamma)dv - [g_w(\underline{v}_s + (1-\gamma)x,\gamma)(\underline{v}_s + (1-\gamma)x)x] - [g_w(\underline{v}_w + \gamma x,\gamma)(\underline{v}_w + \gamma x)x] \quad (5.13)$$

Next, from the second term in (5.6) we get

$$\begin{aligned}
& \int\limits_{\underline{v}_s+(1-\gamma)x}^{\overline{v}_w+\gamma x} \left[v\left(-G_s'(v,\gamma)\right) g_w(v,\gamma) + v(1-G_s(v,\gamma))g_w'(v,\gamma)\right] dv \\
+ \quad & g_w(\underline{v}_s+(1-\gamma)x,\gamma)(\underline{v}_s+(1-\gamma)x)(1-G_s(\underline{v}_s+(1-\gamma)x,\gamma))x \\
+ \quad & g_w(\overline{v}_w+\gamma x,\gamma)(\overline{v}_w+\gamma x)(1-G_s(\overline{v}_w+\gamma x,\gamma))x. \qquad (5.14)
\end{aligned}$$

Equivalently, applying the Leibniz rule to the third term yields

$$\begin{aligned}
& \int\limits_{\underline{v}_s+(1-\gamma)x}^{\overline{v}_w+\gamma x} \left[v\left(-G_w'(v,\gamma)\right) g_s(v,\gamma) + v(1-G_w(v,\gamma))g_s'(v,\gamma)\right] dv \\
+ \quad & g_s(\underline{v}_s+(1-\gamma)x,\gamma)(\underline{v}_s+(1-\gamma)x)(1-G_w(\underline{v}_s+(1-\gamma)x,\gamma))x \\
+ \quad & g_s(\overline{v}_w+\gamma x,\gamma)(\overline{v}_w+\gamma x)(1-G_w(\overline{v}_w+\gamma x,\gamma))x. \qquad (5.15)
\end{aligned}$$

By definition, we have that $G_w(\overline{v}_w+\gamma x,\gamma) = 1 \; \forall \; \gamma$ and $G_s(\underline{v}_s+(1-\gamma)x,\gamma) = 0 \; \forall \; \gamma$. Using this fact and adding up equations (A) through (C) yields (5.7). To save on space, we omit the second-order condition at this stage. ■

5.6.1.3 Optimality condition for uniform distributions (Section 5.3)

If the valuations are uniformly distributed, the objective function for the seller simplifies to

$$\max_{\gamma} \; v_{(2)}^{w,s} = \int\limits_{\underline{v}_w+\gamma x}^{\underline{v}_s+(1-\gamma)x} \frac{v}{\overline{v}_w-\underline{v}_w} dv + \int\limits_{\underline{v}_s+(1-\gamma)x}^{\overline{v}_w+\gamma x} v\left[\left(1-\frac{v-\underline{v}_s-(1-\gamma)x}{\overline{v}_s-\underline{v}_s}\right)\frac{1}{\overline{v}_w-\underline{v}_w} + \left(1-\frac{v-(\underline{v}_w-\gamma x)}{\overline{v}_w-\underline{v}_w}\right)\frac{1}{\overline{v}_s-\underline{v}_s}\right] dv. \qquad (5.16)$$

Taking the derivative of the above expression with respect to γ, equating to zero, and rearranging yields the following optimality condition:

$$x\frac{4x^2\gamma^2 + 4(b-c-x)x\gamma + x^2 + 2(\underline{v}_s-\overline{v}_w)x + \underline{v}_s^2 + \overline{v}_w^2 + \underline{v}_w(\overline{v}_s-\underline{v}_s) - \overline{v}_w(\underline{v}_s+\overline{v}_s)}{(\overline{v}_s-\underline{v}_s)(\overline{v}_w-\underline{v}_w)} = 0.$$

Solving for γ determines the optimal allocation, which is given at

$$\gamma = \frac{1}{2} + \frac{\underline{v}_s - \overline{v}_w + \sqrt{(\overline{v}_s-\underline{v}_s)\cdot(\overline{v}_w-\underline{v}_w)}}{2x}.$$

Finally, we have to verify that we found a local maximum and the second order condition of (5.16) is fulfilled, i.e.

$$-\frac{2}{3}\cdot\frac{6x^2\left[(\overline{v}_w+\gamma x)-(\underline{v}_s+(1-\gamma)x)\right]}{(b-a)(d-c)}<0.$$

Note that above expression is negative, if the bracketed term in the numerator is positive,

$$(\overline{v}_w+\gamma x)-(\underline{v}_s+(1-\gamma)x)>0 \quad\Leftrightarrow\quad \overline{v}_w-\underline{v}_s>(1-2\gamma)x.$$

Substituting (5.8) for γ yields

$$\begin{aligned}
\overline{v}_w-\underline{v}_s &> \left(1-2\left(\frac{1}{2}+\frac{\underline{v}_s-\overline{v}_w+\sqrt{(\overline{v}_s-\underline{v}_s)(\overline{v}_w-\underline{v}_w)}}{2x}\right)\right)x\\
\Leftrightarrow \overline{v}_w-\underline{v}_s &> -\underline{v}_s+\overline{v}_w-\sqrt{(\overline{v}_s-\underline{v}_s)(\overline{v}_w-\underline{v}_w)}\\
\Leftrightarrow \qquad 0 &> -\sqrt{(\overline{v}_s-\underline{v}_s)(\overline{v}_w-\underline{v}_w)}.
\end{aligned}$$

Hence, the second-order condition is indeed non-positive. ■

5.6.1.4 Proposition 4

We begin by proving part (i). Since bidders are symmetric, denote $\underline{v}_s=\underline{v}_w\equiv\underline{v}$ and $\overline{v}_s=\overline{v}_w\equiv\overline{v}$. Thus, $S_w=S_s=\overline{v}-\underline{v}$. Substituting into (5.11) yields

$$\begin{aligned}
\gamma &= \frac{1}{2}+\frac{1}{32x^3}\left(8x^2(\underline{v}-\overline{v})-\lambda(\overline{v}-\underline{v})^2+\sqrt{(\overline{v}-\underline{v})^2\left(\lambda^2(\overline{v}-\underline{v})^2+16\lambda x^2(\overline{v}-\underline{v})+64x^4\right)}\right)\\
&= \frac{1}{2}+\frac{1}{32x^3}\left(8x^2(\underline{v}-\overline{v})-\lambda(\overline{v}-\underline{v})^2+(\overline{v}-\underline{v})\sqrt{\left(\lambda(\overline{v}-\underline{v})-8x^2\right)^2}\right)\\
&= \frac{1}{2}+\frac{1}{32x^3}\left(8x^2(\underline{v}-\overline{v})+\lambda(\overline{v}-\underline{v})^2-\lambda(\overline{v}-\underline{v})^2+8x^2(\overline{v}-\underline{v})\right)=\frac{1}{2}.
\end{aligned}$$

Hence, in absence of asymmetry the second term on the RHS cancels out and the optimal standard is indeed given at $\gamma=\frac{1}{2}$.

For part (ii), we need to show that for any λ, the optimal standard converges to $\gamma=1$ if x becomes small. The optimality condition (5.11) is a complicated equation in x but we can characterize its asymptotic behavior by concentrating only on the higher-order terms. Let $O(\cdot)$ denote the order. If x becomes large, then the polynomial in the second term on the

RHS is of order $O(x^3)$, while all the other lower-order terms can be disregarded. Hence, for $x \to 0$ the term converges to infinity. However, γ is bounded above by 1, yielding $\lim_{x \to 0} \gamma = 1$ for any λ.

It remains to show part (iii) of the proposition. Similarly as above, for $x \to \infty$ the above term converges to zero, as the denominator becomes arbitrarily large. Hence, we have that $\lim_{x\to\infty} \gamma = \frac{1}{2}$. Finally, taking limits with respect of λ while concentrating only the highest-order, $O(\lambda)$, yields $\lim_{x\to\infty} = \frac{1}{2} + \text{sign}[-(S_w \cdot S_s) + (S_w \cdot S_s)] \cdot \infty = \frac{1}{2}$, which confirms our initial conjecture. ■

5.6.2 Simulations with Truncated Normal Distributions

Uniform distributions prove useful due to their mathematical traceability, but it is worthwhile to reexamine the models for a more general class of distributional assumptions. Therefore, we conduct a numerical simulation for the case when bidder $i's$ valuation is a draw from a normal distribution with mean $v_i \sim N(\mu_i, \sigma_i)$. The PDF and CDF for a normal distribution are given by

$$f_i(v) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{v-\mu_i}{\sigma_i})^2} \qquad \text{and} \qquad F_i(v) = \frac{1}{\sigma_i\sqrt{2\pi}} \int_{-\infty}^{v} e^{-\frac{1}{2}(\frac{t-\mu_i}{\sigma_i})^2} dt.$$

Since the valuations are constrained to the closed interval $[\underline{v}_i, \overline{v}_i]$ for $i \in w, s$, it is necessary to truncate the normal distributions both above and below. Let $h_i(v)$ denote the resulting PDF, and $H_i(v)$ the CDF conditional on v_i lying within the considered support, respectively.

$$h_i(v) = \begin{cases} \frac{f_i(v)}{F_i(\overline{v}_i) - F_i(\underline{v}_i)} & \text{if} \quad \underline{v}_i \leq v \leq \overline{v}_i \\ 0 & \text{otherwise} \end{cases} \qquad \text{and} \qquad H_i(v) = \begin{cases} 0 & \text{if} \quad v \leq \underline{v}_i \\ \frac{F_i(v) - F_i(\underline{v}_i)}{F_i(\overline{v}_i) - F_i(\underline{v}_i)} & \text{if} \quad \underline{v}_i < v < \overline{v}_i \\ 1 & \text{if} \quad v \geq \overline{v}_i \end{cases}$$

If bidders' valuations are drawn from normal distributions, we can relax Assumption 2 in the sense that the bidders can now share identical supports. However, to ensure that s ex-ante is still the "strong" bidder in terms of first-order-stochastic dominance, we need to impose some restrictions on the relation of the means and the standard deviations.[32]

[32] In addition, the standard deviations must be sufficiently large relative to the length of the supports. Particularly, a property of the normal distribution is that about 95% of the time a value within two standard deviations around the mean will be realized. If the standard deviation becomes sufficiently small, so does the uncertainty of the seller about the bidders valuations.

Assumption 3 *Let the distributions characterizing the two bidders be such that (i)* $\mu_w \leq \mu_s$*, (ii)* $\sigma_w \geq \sigma_s$*, and (iii)* $\underline{v}_w \leq \underline{v}_s < \overline{v}_w \leq \overline{v}_s$*.*

In what follows, we assume that at least one condition of Assumption 3 is always satisfied with strict inequality. Using this specification, the remainder of this section reconsiders the two scenarios discussed above.

5.6.2.1 Scenario I revisited

Analogously to above, the auctioneer's action in the interim stage will alter the supports of the value distributions. More precisely, in addition to the truncation bounds, also the mean of the value distribution will be shifted upwards yielding $\tilde{\mu}_w(\gamma) = \mu_w + \gamma x$ for the weak, and $\tilde{\mu}_s(\gamma) = \mu_s + (1-\gamma)x$ for the strong bidder, respectively. Let

$$g_w(v,\gamma) = \begin{cases} \frac{f_w(v)}{F_w(\overline{v}_w+\gamma x)-F_w(\underline{v}_w+\gamma x)} & \text{if } \underline{v}_w + \gamma x \leq v \leq \overline{v}_w + \gamma x \\ 0 & \text{otherwise} \end{cases} \quad \text{and}$$

$$g_s(v,\gamma) = \begin{cases} \frac{f_s(v)}{F_s(\overline{v}_s+(1-\gamma)x)-F_s(\underline{v}_s+(1-\gamma)x)} & \text{if } \underline{v}_s + (1-\gamma)x \leq v \leq \overline{v}_w + (1-\gamma)x \\ 0 & \text{otherwise} \end{cases}$$

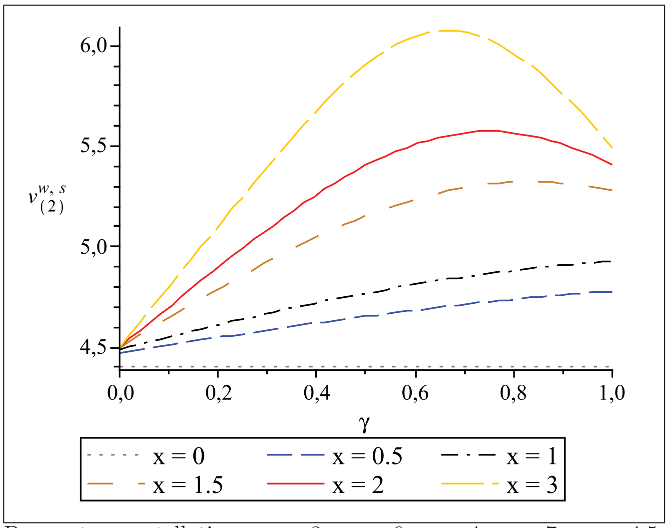
denote the resulting densities for bidder w and s, respectively. The corresponding cumulative distribution functions are stated below.

$$G_w(v,\gamma) = \begin{cases} 0 & \text{if } v \leq \underline{v}_w + \gamma x \\ \frac{F_w(v)-F_w(\underline{v}_w+\gamma x)}{F_w(\overline{v}_w+\gamma x)-F_w(\underline{v}_w+\gamma x)} & \text{if } \underline{v}_w + \gamma x < v < \overline{v}_w + \gamma x \\ 1 & \text{if } v \geq \overline{v}_w + \gamma x \end{cases} \quad \text{and}$$

$$G_s(v,\gamma) = \begin{cases} 0 & \text{if } v \leq \underline{v}_w + (1-\gamma)x \\ \frac{F_s(v)-F_s(\underline{v}_s+(1-\gamma)x)}{F_w(\overline{v}_s+(1-\gamma)x)-F_w(\underline{v}_s+(1-\gamma)x)} & \text{if } \underline{v}_w + (1-\gamma)x < v < \overline{v}_i + (1-\gamma)x \\ 1 & \text{if } v \geq \overline{v}_w + (1-\gamma)x \end{cases}$$

The objective function for the auctioneer is again described by (5.6). However, for expositional simplicity, we conduct a numerical calibration instead of presenting an analytical solution to the seller's problem.[33] Figure 5.7 illustrates the relation between the expected revenue $v_{(2)}^{w,s}$ and the share γ resulting from the simulation for different "cake" sizes x. The parameters can be chosen arbitrarily as to satisfy Assumption 3.

Figure 5.7: Optimal Allocation (Truncated Normal)



Parameter constellation: $\underline{v}_w = 3$, $\overline{v}_w = 6$, $\underline{v}_s = 4$, $\overline{v}_s = 7$, $\mu_w = 4.5$, $\mu_s = 5.5$ and $\sigma_w = \sigma_s = 1$.

For simplicity, let the boundaries of the supports and means be specified as in the previous simulation in Section 5.3, i.e. $\mu_i = \frac{\underline{v}_i + \overline{v}_i}{2}$. In addition, consider the special case where $\sigma_w = \sigma_s = 1$. At the assumed parameter constellation, the degree of asymmetry is thus reflected by the difference of the original means, i.e. $\mu_s - \mu_w = 1$. If $x \leq 1$, Figure 5.7 reveals that it is optimal for the auctioneer to allocate the full cake to the weak bidder. For $x > 1$, expected revenues from the auction are maximized if the bidders are made virtually symmetric, i.e. $\mu_s(v, \gamma) = \mu_w(v, \gamma)$ or $\gamma = \frac{1}{2} + \frac{\overline{v}_s + \underline{v}_s - (\overline{v}_w + \underline{v}_w)}{4x}$.[34] Qualitatively similar findings arise for a large range of parameters satisfying at least one condition of Assumption 3. We next turn to case, where the effect of the seller's action has negative externalities on one of the bidders.

[33]It can be shown that a solution exists, but since this involves the "error function" that is encountered in integrating the normal distribution, a numerical approach appears more promising to infer intuitive insights.

[34]For the given parameters the optimal allocations in the example are thus given by $\gamma_{(x=0.5)} = \gamma_{(x=1)} = 1$, $\gamma_{(x=1.5)} = 0.83$, $\gamma_{(x=2)} = 0.75$ and $\gamma_{(x=3)} = 0.67$.

5.6.2.2 Scenario II revisited

For bidders with diverging tastes, the PDFs and CDFs of the truncated normal distribution are given by

$$g_w(v,\gamma) = \begin{cases} \frac{f_w(v)}{F_w(\overline{v}_w+(2\gamma-1)x)-F_w(\underline{v}_w+(2\gamma-1)x)} & \text{if } \ \underline{v}_w + (2\gamma-1)x \leq v \leq \overline{v}_w + (2\gamma-1)x \\ 0 & \text{otherwise} \end{cases},$$

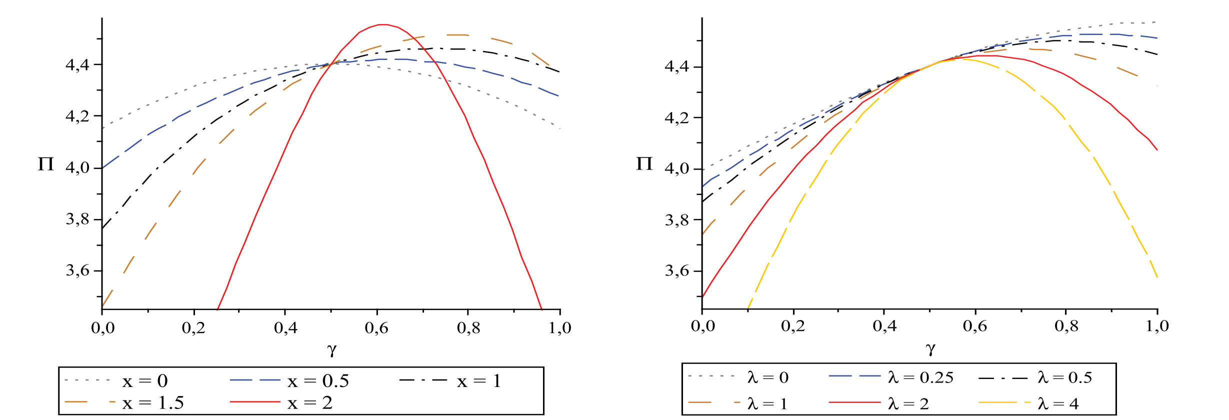
$$g_s(v,\gamma) = \begin{cases} \frac{f_s(v)}{F_s(\overline{v}_s+(1-2\gamma)x)-F_s(\underline{v}_s+(1-2\gamma)x)} & \text{if } \ \underline{v}_s + (1-2\gamma)x \leq v \leq \overline{v}_s + (1-2\gamma)x \\ 0 & \text{otherwise} \end{cases},$$

$$G_w(v,\gamma) = \begin{cases} 0 & \text{if } \ v \leq \underline{v}_w + (2\gamma-1)x \\ \frac{F_w(v)-F_w(\underline{v}_w+(2\gamma-1)x)}{F_w(\overline{v}_w+(2\gamma-1)x)-F_w(\underline{v}_w+(2\gamma-1)x)} & \text{if } \ \underline{v}_w + (2\gamma-1)x < v < \overline{v}_w + (2\gamma-1)x \\ 1 & \text{if } \ v \geq \overline{v}_w + (2\gamma-1)x \end{cases},$$

$$G_s(v,\gamma) = \begin{cases} 0 & \text{if } \ v \leq \underline{v}_w + (1-2\gamma)x \\ \frac{F_s(v)-F_s(\underline{v}_s+(1-2\gamma)x)}{F_w(\overline{v}_s+(1-2\gamma)x)-F_w(\underline{v}_s+(1-2\gamma)x)} & \text{if } \ \underline{v}_w + (1-2\gamma)x < v < \overline{v}_i + (1-2\gamma)x \\ 1 & \text{if } \ v \geq \overline{v}_w + (1-2\gamma)x \end{cases}.$$

The optimality condition is obtained by substituting the above functions into the auctioneer's objective (5.9) and maximizing with respect to γ. Accordingly, Figure 5.8 illustrates the results from a numerical simulation, which we based on the same set of parameters as in Section 5.4. Clearly, the results remarkably resemble those of Figure 5.5. Thus, we conclude that our conjectures from the previous sections are robust to a more general class of distribution functions.

Figure 5.8: Optimal Standard (Truncated Normal)



(a) Revenue in relation to importance of standard (b) Revenue in relation to implementation cost

Parameter constellation: $\underline{v}_w = 3$, $\overline{v}_w = 6$, $\underline{v}_s = 4$, $\overline{v}_s = 7$, $\mu_w = 4.5$, $\mu_s = 5.5$, $\sigma_w = \sigma_s = 1$, $\lambda = 1$ (Figure a), and $x = 0.5$ (Figure b).

References

Ajalin, P., Granö, T., and Nyberg, K. (2004). "Betting on virtually simulated games - Case Hattrick." In "New Business in Computer-Mediated Communities, HUT Software Business and Engineering Institute Technical Reports, HUT-SoberIT-C7.", Sarvas, R. and M. Turpeinen and S. Jokela.

Akerlof, G. A. (1970). "The Market for "Lemons": Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics*, Vol. 84(3), pp. 488–500.

Ariely, D., Ockenfels, A., and Roth, A. E. (2005). "An Experimental Analysis of Ending Rules in Internet Auctions." *Rand Journal of Economics*, Vol. 36(4), pp. 891–908.

Ariely, D. and Simonson, I. (2003). "Buying, Bidding, Playing, or Competing? Value Assessment and Decision Dynamics in Online Auctions." *Journal of Consumer Psychology*, Vol. 13, pp. 113–123.

Arkes, H. R. and Blumer, C. (1985). "The psychology of sunk cost." *Organizational Behavior and Human Decision Processes*, Vol. 35(1), pp. 124–140.

Athey, S. and Haile, P. A. (2002). "Indentification of Standard Auction Models." *Econometrica*, Vol. 70(6), pp. 2107–2140.

Bajari, P. and Hortacsu, A. (2003). "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions." *RAND Journal of Economics*, Vol. 34(2), pp. 329–355.

Bajari, P. and Hortacsu, A. (2004). "Economic Insights from Internet Auctions: A Survey." *Journal of Economic Literature*, Vol. 42(2), pp. 457–486.

Bazerman, M. (1986). *Judgment in Managerial Decision Making*. John Wiley & Sons, Inc.; 2nd ed. Chapter 4.

Benartzi, S. and Thaler, R. (2007). "Heuristics and Biases in Retirement Savings Behavior." *Journal of Economic Perspectives*, Vol. 21(3)(3), pp. 81–104.

Bernard, J.-T., Boldic, D., and Hardy, A. (2002). "The Costs of Natural Gas Distribution Pipelines: The Case of (SCGM), (Quebec)." *Energy Economics*, Vol. 24(5), pp. 425–438.

Bester, H. and Strausz, R. (2000). "Imperfect Commitment and the Revelation Principle: The Multi-Agent Case." *Economics Letters*, Vol. 69, pp. 165–171.

Bhattacharya, U., Holden, C. W., and Jacobsen, S. E. (2008). "Penny Wise, Dollar Foolish: The Left-Digit Effect in Security Trading." *Working Paper.*

Brenner, G. and Brenner, R. (1982). "Memory and Markets, or Why Are You Paying $2.99 for a Widget?" *The Journal of Business*, Vol. 55, pp. 147–158.

Bulow, J., Huang, M., and Klemperer, P. (1999). "Toeholds and Takeovers." *The Journal of Political Economy*, Vol. 107(3), pp. 427–454.

Bulow, J. and Roberts, J. (1989). "The Simple Economics of Optimal Auctions." *Journal of Political Economy*, Vol. 97(5), pp. 1060–1090.

Cantillon, E. (2000). "The Effect of Bidders' Asymmetries on Expected Revenue in Auctions." *Working Paper.*

Cantillon, E. (2008). "The Effect of Bidders' Asymmetries on Expected Revenue in Auctions." *Games and Economic Behavior*, Vol. 62, pp. 1–25.

Carmon, Z. and Ariely, D. (2000). "Focusing on the Forgone: How Value Can Appear So Different to Buyers and Sellers." *Journal of Consumer Research: An Interdisciplinary Quarterly*, Vol. 27(3), pp. 360–370.

Casey, J. T. (1995). "Predicting buyer-seller pricing disparities." *Manage. Sci.*, Vol. 41(6), pp. 979–999.

Castronova, E. (2008). "A Test of the Law of Demand in a Virtual World: Exploring the Petri Dish Approach to Social Science." *CESifo Working Paper No. 2355.*

Cooper, D. and Fang, H. (2006). "Understanding Overbidding in Second Price Auctions: An Experimental Study." *Cowles Foundation Discussion Paper No. 155.7.*

Corns, A. and Schotter, A. (1999). "Can Affirmative Action Be Cost Effective? An Experimental Examination of Price-Preference Auctions." *The American Economic Review*, Vol. 89(1), pp. 291–305.

Ellison, G. (2005). "A Model of Add-on Pricing." *The Quarterly Journal of Economics*, Vol. 120(2), pp. 585–637.

Ellison, G. and Fisher-Ellison, S. (2009). "Search, Obfuscation, and Price Elasticities on the Internet." *Econometrica*, Vol. 77(2), pp. 427–452.

Englmaier, F. and Schmöller, A. (2009a). "Determinants and Effects of Reserve Prices in HATTRICK Auctions." *Working Paper*.

Englmaier, F. and Schmöller, A. (2009b). "Does Bidding for Complex Goods Reflect all Relevant Information? Field Evidence From Online Gaming." *Working Paper*.

Englmaier, F. and Schmöller, A. (2009c). "Efficient Use of Infomation and the Evaluation of Complex Goods - Field Evidence From Online Car Sales." *Working Paper*.

Eso, P. and Szentes, B. (2007). "Optimal Information Disclosure in Auctions and The Handicap Auction." *Review of Economic Studies*, Vol. 74, pp. 705–731.

Gabaix, X. and Laibson, D. (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, Vol. 121(2), pp. 505–540.

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). "Information Acquisition: Experimental Analysis of a Boundedly Rational Model." *American Economic Review*, Vol. 96 (4), pp. 1043–1068.

Genesove, D. and Mayer, C. (2001). "Loss Aversion And Seller Behavior: Evidence From The Housing Market." *The Quarterly Journal of Economics*, Vol. 116(4), pp. 1233–1260.

Gilbert, T., Kogan, S., Lochstoer, L. A., and Ozyildirim, A. (2008). "Investor Inattention and the Market Impact of Summary Statistics." *Working Paper*.

Goeree, J. K. and Offerman, T. (2004). "The Amsterdam Auction." *Econometrica*, Vol. 72(1), pp. 281–294.

Grubbs, F. E. (1969). "Procedures for Detecting Outlying Observations in Samples." *Technometrics*, Vol. 11(1), pp. 1–21.

Hanemann, W. M. (1991). "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review*, Vol. 81(3), pp. 635–647.

Harris, L. (1991). "Stock Price Clustering and Discreteness." *Review of Financial Studies*, Vol. 4(3), pp. 389–415.

Hoppe, T. and Sadrieh, A. (2007). "An Experimental Assessment of Confederate Reserve Price Bids in Online Auction." *Working Paper*.

Hossain, T., Brown, J., and Morgan, J. (2007). "Shrouded Attributes and Information Suppression: Evidence from Field Experiments." *Competition Policy Center. Paper CPC07-075.*

Hotz, J. and Xiao, M. (2007). "Strategic Information Disclosure: The Case of Multi-Attribute Products with Heterogeneous Consumers." *NBER Working Paper No. W11937.*

Kahneman, D., Knetsch, J., and Thaler, R. (1991). "The Endowment Effect, Loss Aversion, and Status Quo Bias." *Journal of Economic Perspectives*, Vol. 5(1), pp. 193–206.

Kahneman, D. and Tversky, A. (1979). "Prospect theory: An analysis of decisions under risk." *Econometrica*, Vol. 47, pp. 313–327.

Kamins, M. A., Dreze, X., and Folkes, V. S. (2004). "Effects of Seller-Supplied Prices on Buyers' Product Evaluations: Reference Prices in an Internet Auction Context." *Journal of Consumer Research: An Interdisciplinary Quarterly*, Vol. 30(4), pp. 622–628.

Kirkegaard, R. and Overgaard, P. B. (2005). "Pre-Auction Offers in Asymmetric First-Price and Second-Price Auctions." *CIE Discussion Paper.*

Klemperer, P. D. (1998). "Auctions With Almost Common Values: The 'Wallet Game' and its Applications." *European Economic Review*, Vol. 42(3-5), pp. 757–769.

Klemperer, P. D. (Editor) (2000). *The Economic Theory of Auctions.* Edward Elgar; 1st edition.

Knetsch, J. L. (1989). "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *American Economic Review*, Vol. 79(5), pp. 1277–1284.

Krishna, V. (2002). *Auction Theory.* Academic Press; 1st edition.

Lee, Y. H. and Malmendier, U. (2007). "The Bidder's Curse." *NBER Working Paper No. 13699.*

Levin, D. and Smith, J. L. (1994). "Equilibrium in Auctions with Entry." *American Economic Review*, Vol. 84(3), pp. 585–599.

Levin, D. and Smith, J. L. (1996). "Optimal Reservation Prices in Auctions." *The Economic Journal*, 106(438), pp. 1271–1283.

Lucking-Reiley, D. (1999). "Using Field Experiments to Test Equivalence between Auction Formats: Magic on the Internet." *American Economic Review*, Vol. 89(5), pp. 1063–1080.

Lucking-Reiley, D. (2000). "Auctions on the Internet: What's Being Auctioned, and How?" *Journal of Industrial Economics*, Vol. 48, pp. 227–252.

Lucking-Reiley, D., Bryan, D., Prasad, N., and Reeves, D. (2007). "Pennies from Ebay: The Determinants of Price in Online Auctions." *Journal of Industrial Economics*, Vol. 55(2), pp. 223–233.

MacKinnon, J. and White, H. (1985). "Some heteroskedasticity consistent covariance matrix estimators with improved finite sample properties." *Journal of Econometrics*, Vol. 29, pp. 53–57.

Manzini, P. and Mariotti, M. (2006). "Two-Stage Boundedly Rational Choice Procedures: Theory and Experimental Evidence." *IZA Discussion Papers No. 2341.*

Manzini, P. and Mariotti, M. (2007). "Sequentially Rationalizable Choice." *American Economic Review*, Vol. 97(5), pp. 1824–1839.

Maskin, E. and Riley, J. (2000). "Asymmetric Auctions." *Review of Economic Studies*, Vol. 67, pp. 439–545.

McAfee, R. P. and McMillan, J. (1987). "Auctions and Bidding." *Journal of Economic Literature*, Vol. 25, pp. 699–738.

Myerson, R. (1981). "Optimal auction design." *Mathematics of Operations Research*, Vol. 6, pp. 58–63.

Nicklisch, A. and Salz, T. (2008). "Reciprocity and status in a virtual field experiment." *Working Paper.*

Niederhoffer, V. (1965). "Clustering of Stock Prices." *Operations Research*, Vol.13(2), pp. 258–265.

Paarsch, H. J. and Hong, H. (2006). *An Introduction to the Structural Econometrics of Auction Data.* MIT Press.

Plott, C. R. and Zeiler, K. (2005). "The Willingness to PayûWillingness to Accept Gap, the "Endowment Effect," Subject Misconceptions, and Experimental Procedures for Eliciting Valuations." *American Economic Review*, Vol. 95(3), pp. 530–545.

Plott, C. R. and Zeiler, K. (2007). "Exchange Asymmetries Incorrectly Interpreted as Evidence of Endowment Effect Theory and Prospect Theory?" *American Economic Review*, Vol. 97(4), pp. 1449–1466.

Pope, D. G. (2008). "Reacting to Rankings: Evidence from "America's Best Hospitals"." *Working Paper.*

Reiley, D. (2006). "Field Experiments on the Effects of Reserve Prices in Auctions: More Magic on the Internet." *RAND Journal of Economics*, Vol. 37(1), pp. 195–211.

Riley, J. and Samuleson, W. (1981). "Optimal auctions." *The American Economic Review*, Vol. 71, pp. 381–392.

Rosenkranz, S. and Schmitz, P. W. (2007). "Reserve Prices in Auctions as Reference Points." *Economic Journal*, Vol. 117(520), pp. 637–653.

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." *The American Economic Review*, Vol. 92(4), pp. 1093–1103.

Rothkopf, M., Harstad, M., and Fu, Y. (2003). "Is Subsidizing Inefficient Bidders Actually Costly?" *Management Science*, Vol. 49(1), pp. 71–84.

Schwarz, M. and Sonin, K. (2005). "The Variable Value Environment: Auctions and Actions." *CEFIR Working Paper W0020.*

Shogren, J. F., Shin, S. Y., Hayes, D. J., and Kliebenstein, J. B. (1994). "Resolving Differences in Willingness to Pay and Willingness to Accept." *The American Economic Review*, Vol. 84(1), pp. 255–270.

Simon, H. A. (1955). "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics*, Vol. 69(1), pp. 99–118.

Simon, H. A. (1978). "Rationality as a Process and Product of Thought." *American Economic Review*, Vol. 68(2), pp. 1–16.

Simonsohn, U., Karlsson, N., Loewenstein, G., and Ariely, D. (2008). "The tree of experience in the forest of information: Overweighing experienced relative to observed information." *Games and Economic Behavior*, Vol. 62(1), pp. 263–286.

Sonnemans, J. (2006). "Price Clustering and Natural Ressistance Points in the Dutch Stock Market: A Natural Experiment." *European Economic Review*, Vol. 50, pp. 1937–1950.

Staw, B. M. (1976). "Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action." *Organizational Behavior and Human Performance*, Vol. 16, pp. 27–44.

Thaler, R. (1980). "Towards a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization*, Vol. 1, pp. 39–60.

Trautmann, S. T. and Traxler, C. (2009). "Reserve prices as reference points ù Evidence from auctions for football players at hattrick.org." *Working Paper.*

Vickrey, W. (1961). "Counterspeculation and Competitive Sealed Tenders." *Journal of Finance*, Vol. 16(1), pp. 8–37.

Yepez, R. A. (2008). "A Cost Function for the Natural Gas Transmission Industry." *Engineering Economist*, Vol. 53(1), pp. 68–83.

Zwick, R., Rapoport, A., Lo, A. K. C., and Muthukrishnan, A. V. (2003). "Consumer Sequential Search: Not Enough or Too Much?" *Marketing Science*, Vol. 22(4), pp. 503–519.

List of Figures

List of Tables

Curriculum Vitae

10/2005 - present	Research and Teaching Assistant, Ph.D. Program in Economics Munich Graduate School of Economics Ludwig-Maximilians-Universität, Munich, Germany
10/2002 - 07/2003	Studies in Economics, Visting Student University of Southampton, England
10/2000 - 07/2005	Studies in Economics, Diplom-Volkswirt Ludwig-Maximilians-Universität, Munich, Germany
09/1999 - 06/2000	Military Service, Murnau, Germany
07/1999	Abitur Gymnasium Bad Tölz, Germany
07/21/1979	Born in Tegernsee, Germany