ESSAYS ON EMPIRICAL ANALYSIS OF CONTINUOUS-TIME MODELS OF INDUSTRIAL ORGANIZATION

by

Denis Nekipelov

Department of Economics
Duke University

Date: ______________________________

Approved:

____
Han Hong, Supervisor

____
Shakeeb Khan, Supervisor

____
Patrick Bajari

____
Curtis Taylor

____
Christopher Timmins

____
Paul Ellickson

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Economics
in the Graduate School of
Duke University

2008

ABSTRACT

ESSAYS ON EMPIRICAL ANALYSIS OF CONTINUOUS-TIME MODELS OF INDUSTRIAL ORGANIZATION

by

Denis Nekipelov

Department of Economics
Duke University

Date: ______________________

Approved:

____
Han Hong, Supervisor

____
Shakeeb Khan, Supervisor

____
Patrick Bajari

____
Curtis Taylor

____
Christopher Timmins

____
Paul Ellickson

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics in the Graduate School of Duke University

2008

Abstract

The dissertation consists of three essays. The first essay describes and estimates the model of bidding on eBay. Internet auctions (such as eBay) differ from the traditional auction format in that participants 1) typically face a choice over several simultaneous auctions and 2) often have limited information about rival bidders. Since existing economic models do not account for these features of the bidding environment, it should not be surprising that even casual empiricism reveals a sharp discrepancy between the predictions of existing theory and the actual behavior of bidders. In this paper, I show that the presence of multiple, contemporaneous auctions for similar items coupled with uncertainty regarding rival entry can explain both features. I analyze these features in a continuous-time stochastic auction model with endogenous entry, in which bidder types are differentiated by their initial information regarding the entry process. Empirical estimates using eBay auctions of pop-music CDs confirm my theoretical prediction that the rate of entry depends on price. I then test my model against alternative explanations of observed bidding behavior using a detailed field experiment.

The second essay is on empirical analysis of executive compensation in the continuous-time environment. In this essay, I develop a methodology for the identification and non-parametric estimation of a continuous-time principal-agent model. My framework extends the existing literature on optimal dynamic contracts by allowing for the presence of unobserved state variables. To accommodate such heterogeneity, I develop an estimation method based on numerically solving for the optimal non-linear manager's response to the restrictions of the contract. To demonstrate this feature, I apply my methodology to executive contracts from the retail apparel industry.

The third essay provides a tractable methodology for the construction and structural estimation of continuous time dynamic models. The specific class of models covered by my framework includes competitive dynamic games where there are no direct spillovers between objective functions of players. I develop an estimation methodology based on the properties of the equilibrium of the model. The methodology that I design can be applied to welfare and revenue analysis of large dynamic models. As an example, I compute the revenue and welfare gains for a counter-factual exercise in which the eBay auction website changes the format of its auctions from the second price format to a flexible ending format.

Contents

List of Tables

List of Figures

Acknowledgements

I am deeply indebted to Han Hong as well as Pat Bajari, Chris Timmins, Paul Ellickson, and Curt Taylor for their help and continuous encouragement. I would like to thank Peter Arcidiacono, Pat Bayer, Lanier Benkard, Arie Beresteanu, Hanming Fang, Jesus Fernandez-Villaverde, Joseph Hotz, Rachel Kranton, John Lazarev, Jon Levin, Marjorie McElroy, Robert McMillan, Juan Rubio-Ramirez, Andrew Sweeting, and Duncan Thomas for their suggestions and fruitful discussions, all participants of the Microeconometrics and the Microeconomic Theory lunches at Duke, Econometrics Lunch and IO workshop at Stanford, and Structural Lunch at Stanford GSB.

Chapter 1

Entry Deterrence and Learning Prevention on eBay

1.1 Introduction

Since it went online in 1995, eBay has provided both consumers with an efficient market for the exchange of goods and researchers with fertile ground for testing academic theories about auctions. Although empirical evidence from this rich environment has validated many aspects of standard auction theory, in general, auction theory provides little information about optimal strategic behavior in a dynamic multiple-auction model. Many observed features of bidding behavior contradict predictions from the standard auction models, suggesting that these models provide an incomplete picture of online auctions. In particular, it has been widely observed that bidders on eBay tend to bid multiple times for the same object and cluster their bids in the final seconds of a given auction.[1] Such behavior is inconsistent with a standard second-price auction, whereby bidders simply submit their valuations once and for all. This paper instead proposes a dynamic model of competing auctions that can both rationalize the observed behavior and provide additional testable restrictions. These restrictions are then evaluated using data from a novel field experiment.

The existence of many simultaneous auctions is a defining feature of eBay. It is also the primary reason why static auction theory cannot explain observed online bidding patterns. Because similar items are frequently auctioned in close temporal proximity, the number of entrants into a particular auction will depend on how the

[1] See, for example, [BH03].

price in that auction relates to prices in other auctions. The endogenous nature of entry, coupled with bidders' uncertainty regarding the number of rivals they expect to face in a given auction, creates two opposing strategic incentives that are absent from the simple static framework: entry deterrence and learning prevention. On the one hand, should a bidder find an object for which there are no current bids (given a reasonably attractive starting price), she has a strong incentive to bid immediately, thereby deterring the entry of potential rivals. On the other hand, once an auction has been entered, more informed bidders will have an incentive to delay their bids until the final seconds of the auction, in order to prevent their private information from being incorporated into the price. I will show that the tension between these two opposing forces can explain the coexistence of both early and late bidding. The structure of the model also highlights the importance of learning, even in settings where bidder valuations are uncorrelated.

To formalize this intuition, I propose and empirically estimate a continuous-time structural auction model with endogenous entry. I first estimate the structural parameters of the model using data from a single auction category (pop music CDs). I then conduct a controlled experiment aimed at identifying the impact of entry deterrence on optimal bidding behavior. The rationale behind this experiment depends on the fact that the incentive to deter entry depends on the thickness of the market. In thicker markets (which have more auctions per bidder), entry into individual auctions should be significantly more sensitive to price than in markets that are thinner. Therefore, an exogenous increase in the number of objects being auctioned should be associated with a significant increase in early bidding. Consistent with the predictions of the model, the experiment reveals that early bidding is indeed significantly more common when the number of contemporaneous auctions increases.

This paper contributes to both the auction literature and a growing field of research aimed at estimating models of dynamic strategic interaction. The theoretical model illustrates the importance of considering auctions within the context of market interactions with endogenous entry. I show that even if bidder's valuations are uncorrelated, market thickness and the composition of bidder types in the market each have a significant effect on aggressive bidding. In particular, the strategic behavior of individual bidders depends on the characteristics of every available auction, as well as all available information regarding potential rivals. This creates a complex learning process in which bidders must infer the number of potential entrants and their valuations from observations on price. Although solving the full strategic game in a high dimensional space is clearly intractable, by reformulating the problem with a dynamic competitive mechanism and working in continuous time, I am able to reduce the burden significantly. In particular, I assume that instead of reacting to particular rivals, bidders formulate a best-response to the price movement. This structure allows me to reduce the individual bidding problem to one of single-agent optimal control and analyze an equilibrium similar in flavor to the competitive equilibria studied in the stochastic finance literature.[2] The techniques developed here can easily be extended to other dynamic settings. In Chapter 2, I use a similar approach to tackle the dynamic principal-agent model of [HM87].

By modeling entry as a Poisson process with a price-dependent frequency, I am able to capture the key features of endogenous entry without solving the full strategic game. The model employs a dynamic competitive mechanism in which bidders directly influence only the price, which then affects the payoffs of rival bidders through

[2]See, for instance [Coc01], and [Shi99].

their market interaction. Uncertainty is captured as unobserved heterogeneity across auctions using a one-dimensional "visibility" parameter. While bidders have different a priori information regarding the visibility of the auction, they can only learn its value during the auction by observing the price path. Although I assume that the auction visibility is a fixed parameter in a single auction, it is used as a reduced-form characteristic of an equilibrium in a multiple auction environment. Therefore, visibility is endogenous in a general equilibrium context.

Although the bid function in my model does not yield a closed-form solution, I am able to reduce the computational burden significantly by formulating the model in continuous time and exploiting a convenient linear algorithm to solve the model numerically. My estimation method employs a form of indirect inference and contains a method of solving for individual bidding strategies that only requires a fixed number of iterative steps, yielding a much faster procedure than traditional value function iteration. The approach is based on matching the empirically-observed distribution of equilibrium responses to its theoretical counterpart. In an auction, if bidders respond optimally to the movement of price, their bids should replicate the price process over time. This idea is implemented by minimizing the distance between these two distributions. I establish consistency and asymptotic normality of the structural parameter estimates, as well as non-parametric identification of the auction model emphasized here.

To estimate the model, I use a unique dataset of auctions for pop-music CDs on eBay - a collection of auctions over homogeneous items for which bidders arguably have uncorrelated valuations. The estimation results validate the predictions of my theoretical model. Specifically, I observe that the entry rate is a decreasing function of price, while the average entry rate is closely related to the average number of bidders per auction. In addition, the model predicts that prices should adjust more frequently towards the end of the auction, which also coincides with the empirical observations. The structural parameters also allow me to perform a counterfactual analysis to examine how equilibrium bidding responds to the elasticity of entry with respect to price and bidders' beliefs. In particular, I quantify the effect of strategic early bidding in the auctions and find that it contributes more to the observed bidding patterns than non-strategic uninformed bidding.

There have, of course, been other attempts at explaining the observed features of eBay bidding behavior.[3] To distinguish my model from competing theories, I conduct a field experiment that highlights a unique feature of my framework: the relationship between entry deterrence and market thickness. Specifically, according to my model, any event that shifts the rate of entry into a group of auctions should also change optimal bidding strategies. In particular, the incentive to deter entry should depend on the thickness of the market (other things being equal). In thicker markets (holding the number of potential bidders constant), entry should be significantly more sensitive to price, making entry deterrence (and the early bidding associated with it) more common in larger markets. This prediction is not shared by the alternative models listed above. Therefore, I can directly test the validity of my hypothesis against these alternative models via a controlled experiment that artificially increases the size of the market for a specific CD. My empirical results reveal that an expansion in the size of the market indeed increases the frequency of early bidding, confirming my

[3]Existing attempts to explain the eBay data include the tacit collusion proposed in [OR02], technical features of eBay (lack of information about the proxy bidding mechanism) offered in [OR02], multiplicity of listings in [JS02] and [PS04]) and behavioral explanations, such as bidder's irrationality or uncertainty about own private value as in [Ras03] and [Hos04], risk-aversion of bidders in [AHS06] and the possibility of common values in [BH03].

predictions.[4]

The structure of the paper is as follows. Section 2 describes the data used in the structural estimation, and Section 3 describes the model and derives the optimal bidding strategy. Section 4 outlines the identification strategy and explains the estimation procedure. Section 5 contains the main structural estimates of the model obtained using the data from eBay and the results of counterfactual experiments. Section 6 presents the results of the field experiment and a broader discussion. Section 7 concludes.

1.2 Bidding for pop-music CDs on eBay

My theoretical framework focuses on private value auctions. While there is no common-value component, individual bidders have to learn about the attractiveness of competing auctions by observing the price process. To carry out the analysis I will need to identify a class of eBay auctions for which this assumption is appropriate.[5] I have chosen to study music CDs from the recording artist Madonna for which, I will argue below, the private value assumption is appropriate. First, a CD is a relatively inexpensive item, so its purchase should have minor wealth effects. Second, variation in the market price of CDs sold outside of eBay is likely to be small, implying that most bidders are consuming CDs for their intrinsic value, rather than buying them for resale purposes. Finally, focusing on a single recording artist should further limit

[4]A unique feature of my field experiment is that it focuses on both homogeneous and inexpensive items. Therefore, issues specific to common-value items, such as the winner's curse, are unlikely to arise. Most previous field experiments on eBay have focused on common-value items. For instance, [Yin03] studies bidding for computers on eBay, other studies such as [Rei06], [BH03] and [OR02] of eBay discuss bidding for collectible items which also possess common value components.

[5]A private value structure requires that valuations of different bidders be uncorrelated. In practice, I will assume that a private value setting is appropriate so long as these correlations are sufficiently small compared to the variance of valuations.

unobserved heterogeneity in the bidder population and the uncertainty in the quality of individual items.

Data from eBay have been used by economists in several empirical studies, mainly in the context of common values.[6] My dataset has a unique feature that it contains information about both the auctioned objects (including information about the seller) as well as the bidding process itself. The data was collected using a *Perl* "spider" program, which gathered information about both the items and the bidding histories.[7] The information on bidding behavior characterizes the bid profile for each item, along with the bid history, includes an item identifier, bid amounts, and a bidder identifier. The sample statistics (timing of bids and their quantities) for this dataset are contained in Table A.1.[8] In Table A.1, the duration of all auctions is normalized to 1 the reason for which will be explained in a subsequent section). Table A.1 reveals that the average bid is close to $14 and the average bidding time is very close to the end of the auction. The abundance of information in my dataset facilitates the estimation of a structural continuous-time model and a related field experiment.

Along with the bidding profile, I collected additional information about each auction in the dataset. The variables, summarized in Table A.1, include the duration of the auction, the buy-it-now price (if offered), the shipping cost, the condition of the

---

[6] The eBay computer category studied in [Yin03] has an issue with product authenticity. Common value component is essential for bidding behavior in many other empirical studies such as [BH04] and [OR02].

[7] The program was designed to work as follows. First it submitted a search query for music CDs sold through auctions (no eBay stores), and sorted the items according to the auction ending time. The program selects up to 20 pages of items so that the auctions for those items end in at most 4 hours. The program browses individual pages for those items and saves the exact time when the auction ends, the rating and the nickname of the seller, the characteristics of the auction for a CD. Then the program sleeps for 4 hours and goes to the page where eBay saves the bidding history. The process repeats the necessary number of times.

[8] The tables described in this and further sections can be found in Appendix A

CD (used or new), the seller's feedback score, the proportion of the seller's positive feedback, and a dummy variable indicating that the item's page includes a photo of the CD cover. Most sellers in the dataset have very high scores, and positive feedback levels. The average duration of the auction is 6.5 days, reflecting the fact that sellers can avoid additional listing fees by choosing an auction lasting less than 7 days. The buy-it-now price is very close to the average bid and the buy-it-now option is offered approximately half of the time. The condition variable reveals that half of the CDs on the market are used, while the picture dummy indicates that most auctions include a photo of the CD. Approximately 83% of the sellers in my dataset were located in the United States. Other sellers were primarily located in Argentina, Australia, Europe and Japan. I should also note that the four best-selling Madonna albums (according to Amazon.com) accounted for 16% of the eBay market. A "top 4" indicator variable will therefore be used to control for additional heterogeneity across auctions.

In Table A.2, I examined some features of the data that are related to the bidding behavior describing my model. To describe these features, I first constructed two additional variables: a variable corresponding to the arrival rate of bids (i.e. the number of bids in the auction per unit of time) and another variable equal to the size of the price jump (in the beginning of the auction, it is simply the first bid). I analyzed how these variables are affected by the auction duration, the current prices and other characteristics of auctions. The first column contains the output from a fixed effect regression that explains the observed size of bid increments. The estimates reveal that bids become more incremental toward the end of the auction. Moreover, the price tends to jump higher when the price level is high.

In column 2 the dependent variable is the number of bidders who are active over the course of the auction. The estimates from this regression reveal that price is a significant determinant of entry: the rate of entry of bidders is lower in auctions with higher starting prices. These findings imply that ignoring the dependence of bidding behavior on price in eBay auctions, as in a static private value auction model, will lead to incorrect conclusions.

The regression analysis reported in Table A.2 not only confirms that the basic features of the auction data are consistent with my theoretical model but also suggests relevant auction observables that can be used to control for cross-auction heterogeneity in my dataset. These characteristics will be used in the structural econometric model which I develop later in this paper. Column 3 examines the intensity of bidding in the auctions. Here, we can see that the number of bids is lower if starting price is high and higher for the most recent albums, while a higher feedback score (of the seller) increases the number of bids. Similarly, Column 4 shows that the results for the number of bidders are close to the results for the number of bids in Column 3. Therefore, we can conclude that the observed characteristics of cross-auction variation are significant and have sensible coefficients in these simple regressions. For this reason, I will use them as shifters in my structural model.

1.3 Theoretical model

1.3.1 The Setup

eBay is a large Internet auction site where similar items are often available in multiple, simultaneous auctions. eBay bidders don't just choose how much to bid in a given auction, but also which auction to participate in. The choice of auction clearly depends on both the current posted price as well as the relative chance of winning, requiring bidders to form expectations regarding the likely outcome of the bidding process and then update these beliefs as the auction unfolds. In my model I consider independent private value auctions. Learning behavior in my model comes from the learning about the evolution of the future price process rather than the object itself.

In my model I analyze a single auction within a context of a large auction market. It is a continuous time auction model with both continuous actions (bids) and states (prices). Entry into a particular auction is characterized by a stochastic time-heterogeneous Poisson process that determines the number of bidders that enter the auction per unit of time. The arrival rate or frequency of this Poisson process can be thought of as the "instantaneous demand" for the object on auction. Submitted bids cannot be withdrawn; non-participants simply submit bids of zero. Withdrawal of bids is almost never observed in the dataset, the only two incidences of bid withdrawal was due to submission of very large bids by mistake[9]. The price in this auction is described as a continuous-time heterogeneous Poisson process (which is a Markov process). A bidder who arrives in the auction responds optimally to the entire path of price process rather than to actions of particular rival bidders.

In my model the behavior of sellers is taken as exogenously given. In general, it is also possible that sellers strategically choose the time to list their item(s), as well as the relevant auction parameters, including a reserve price (which can be kept secret) and the "buy-it-now" option[10]. I will consider relaxing this assumption in my

[9]Entry and exit from in auctions corresponding to the withdrawal of bids is discussed for the case of a static English auction in [Izm03]. The setting with arrival of new buyers in the bargaining context has been considered in [FS07].

[10]Many eBay sellers choose to list their items as "buy-it-now" only, meaning that we observe both items sold in the auction format and items sold at fixed prices. The formation of such an equilibrium is consistent with the theory of search and matching as examined in [Pis90], [BMSng] and [Lag00], which study search frictions in different markets. In particular, it has been shown that in a costly search environment, a distribution of different prices for the same item can be supported in equilibrium.

future work. However, tackling the full two-sided market with endogenous buyers and sellers is substantially more complex to analyze[11].

The role of bidders' searching behavior is captured in my paper by modeling the entry process directly. The instantaneous demand for the object is clearly a function of both price and time, but it also depends on bidders' beliefs regarding the ultimate outcome of the auction process (i.e. whether they will win and how much they will have to pay). This in turn depends on the overall "visibility" of the auction: how attractive it is to enter. This notion of visibility is reflected in an idiosyncratic characteristic that is specific to each individual auction. Note that this structure captures the two salient features of the eBay environment noted above: multiple listings of identical items and heterogeneity in bidder information. The dependence of instantaneous demand on the price of the object is motivated by the role of search. Specifically, if the entry decision is based on the reservation value from a search process, a bidder will enter into an auction where the price of the object is lower (*ceteris paribus*) with higher probability. This means that the instantaneous demand will be a decreasing function of price (holding all other characteristics of the item constant). This structure suggests that the price process in general depends to a family of Poisson processes parametrized by the visibility. To respond optimally to the price movement bidders will need to identify the visibilty and, thus, the price process per se.

The visibility parameter is intended to capture idiosyncratic differences in entry rates for different auctions of similar items. It is fixed for a specific auction and does

[11]In the dynamic game describing the auction market, the first stage of the game characterizes optimal reserve price setting by sellers and search for an auction by buyers. In the second stage the buyers compete with each other in particular auctions. A similar setup where the second stage is described by a static second-price auction is considered in [PS97].

not change over time. The visibility is closely related to the elasticity of entry into auctions with respect to price. By introducing this parameter, I am able to capture a form of unobserved heterogeneity that is fixed across auctions. Intuitively, the visibility of the auction should reflect heterogeneous effects coming from the supply-side, in particular, the extent of a seller's participation on eBay. For instance, if a seller is selling a variety of items (or perhaps owns an eBay store), some of her items might be cross-listed. As a result, bidders who enter into her auction might follow a link from one of her other items, rather than via a purposeful search, giving her greater visibility. All else equal, cross-listed auctions should have more bidders than auctions with no cross-listings. Another source of heterogeneity of entry coming from the supply-side effects are listing errors by the seller. For example, a seller might misspell the name of the item or make an error in the item's profile. As a result, the item might not appear in the search results by keyword or might appear lower in the queue than it would have had it not been misspelled. Finally, some bidders could simply be loyal; if some sellers have a long history of producing satisfied customers, it's likely that they will have repeat customers. These supply-side forces will all influence the visibility of the auction.

The optimal strategy of the bidders depends on the visibility parameter. Bidders do not observe the exact value of this parameter.[12] They have prior beliefs in the form of probability distributions over its value. They can then update these beliefs (i.e., learn) over the course of the auction. Initial heterogeneity in beliefs is captured by assuming that bidders possess heterogeneous priors, possibly due to differences in their experience or sophistication. It is important to emphasize that an auction's

[12] The bidders can distinguish between more visible and less visible auctions, but they cannot predict the exact quantitative effect for a specific auction.

visibility is not related to the properties of the object itself: uncertainty regarding visibility is present even if every bidder is completely certain about the properties of the object. In this way, the learning concept considered in this paper is different from the typical common value interpretation, in which bidders are learning about the quality of the object on auction. This assumption emphasizes a specific feature of online auctions: even if the bidder is completely certain about the quality of the object itself, she can be uncertain about the group of potential rivals. The bidder will then need to adjust her strategic behavior in response to the distribution of rival bidding behavior.

1.3.2 Motivating example

Consider the following three-period model of an auction with price-dependent entry. The type of the auction (visibility) is determined by a parameter θ. The price each period is determined by the second-highest outstanding bid, except for the first period where it is determined by the starting price set by the seller. The number of bidders entering in period t, if the price of the object in the previous period is p has a Poisson distribution with parameter $\lambda\,(t,\,p,\,\theta)$. Bidders know the functional form of $\lambda(\cdot)$, but do not observe the visibility θ. Bidders who enter into the auction receive normal signals $z = \theta + \sigma\xi$ (beliefs about the visibility) , where ξ is a standard normal random variable. Bidders observe price history up to the current period. The auction will be characterized by the starting price p_0 and the visibility θ. The auction has the following timing.

Period 1

Number of entering bidders is generated from $Poisson\,(\lambda\,(1, p_0, \theta))$, bidders receive signals z_i. Bidders make bids b^1.

Period 2

Bidders observe the price in the first period $p_1 = \max_{i,\, i \neq argmax_j b_j^1} b_i^1$. Number of entering bidders is generated from $Poisson\left(\lambda\left(2, p_1, \theta\right)\right)$ independent from the process determining the number of bidders in the first period, new bidders receive signals z. Bidders make bids b^2.

Period 3

Bidders observe the price in the second period $p_2 = \max_{i,\, i \neq argmax_j \{b_j^2, p_1\}} \{b_i^1, p_1\}$. Number of entering bidders is generated from $Poisson\left(\lambda\left(3, p_2, \theta\right)\right)$, new bidders receive signals z. Bidders make bids b^3. The bidder with the highest bid wins the auction and pays the amount of the second-highest bid.

The type of each bidder i is determined by her valuation $v_i \sim F(v)$ and the initial signal z_i. Bidders do not observe particular bidders who submitted bids. The information set is common among bidders who have not submitted bids and in period 1 it is $\{p_0\}$, in period 2 it is $\{p_1, p_0\}$, and in period 3 it is $\{p_0, p_1, p_2\}$. If the bidder submits bids, then information set will also contain the own bids of the bidder.

Consider a symmetric Bayes-Nash equilibrium in the presented model. Bidding strategy for bidders in period 1 is $b^1(\cdot; v, z) : \{p_0\} \mapsto \mathbb{R}$, in period 2 it is $b^2(\cdot; v, z) : \{p_1, p_0\} \mapsto \mathbb{R}$, and in period 3 it is $b^3(\cdot; v, z) : \{p_0, p_1, p_2\} \mapsto \mathbb{R}$. Now we can write the payoffs for the bidders entering into the auction in each of the three periods.

Period 3 entrants:

Let $f_{p_2}^3\left(p, n_1, n_2, n_3, b_i^3\right)$ be the density of the distribution of the highest bid of $n_1 + n_2 + n_3 - 1$ of bidders other than bidder i given that this bid is below b_i^3 and each winning bid should exceed p_2. For entrants in the third period the payoff can be written as:

$$\pi_i^{(3)}\left(b_i^3; v_i, z_i, p_0, p_1, p_2\right) = \int (v_i - p)\, f_{p_2}^3\left(p, n_1, n_2, n_3, b_i^3\right) \widehat{N}_{z_i,p_0,p_1,p_2}\left(dn_1, dn_2, dn_3\right) dp$$

where $\widehat{N}_{z_i,p_0,p_1,p_2}(\cdot)$ is the Poisson measure reflecting posterior distribution of the number of entering bidders given the type of the bidder and price realizations. Denote the solution to the optimal bidding problem (which is going to be a function of the price path and the type of the bidder) by b_i^{3*}

Period 2 entrants:

Denote $f_{p_1}^2\left(s_1, s_2, n_1, n_2\right)$ the joint distribution of the first and second order statistics for bids of $n_1 + n_2 - 1$ bidders other than bidder i in period 2, given that they should exceed p_1. For entrants in the second period the payoff can be written as:

$$\begin{aligned}
&\pi_i^{(2)}\left(b_i^2; v_i, z_i, p_0, p_1\right) \\
&\qquad = \int\limits_{s_2 > b_i^2} \pi_i^{(3)}\left(b_i^{3*}; v_i, z_i, p_0, p_1, s_2\right) f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) \widehat{N}_{z_i,p_0,p_1}\left(dn_1, dn_2\right) ds_1 ds_2 \\
&\qquad + \int\limits_{\substack{s_2 < b_i^2 \\ s_1 > b_i^2}} \pi_i^{(3)}\left(b_i^{3*}; v_i, z_i, p_0, p_1, b_i^2\right) f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) \widehat{N}_{z_i,p_0,p_1}\left(dn_1, dn_2\right) ds_1 ds_2 \\
&\qquad + \int\limits_{s_1 < b_i^2} \pi_i^{(3)}\left(b_i^{3*}; v_i, z_i, p_0, p_1, s_1\right) f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) \widehat{N}_{z_i,p_0,p_1}\left(dn_1, dn_2\right) ds_1 ds_2.
\end{aligned}$$

Here I denoted the posterior distribution of the number of players in periods 1 and 2 $\widehat{N}_{z_i,p_0,p_1}(\cdot)$.

Period 1 entrants:

Denote $f_{p_0}^1\left(s_1, s_2, n_1\right)$ the joint distribution of the first and second order statistics for bids of $n_1 - 1$ bidders other than bidder i in period 1, given that they should exceed

p_0. For entrants in the first period the payoff can be written as:

$$\pi_i^{(1)}\left(b_i^1; v_i, z_i, p_0\right) = \int\limits_{s_2 > b_i^1} \pi_i^{(2)}\left(b_i^{2*}; v_i, z_i, p_0, s_2\right) f_{p_0}^1\left(s_1, s_2, n_1\right) \widehat{N}_{z_i, p_0}\left(dn_1\right) ds_1 ds_2$$

$$+ \int\limits_{\substack{s_2 < b_i^1 \\ s_1 > b_i^1}} \pi_i^{(2)}\left(b_i^{2*}; v_i, z_i, p_0, b_i^1\right) f_{p_0}^1\left(s_1, s_2, n_1\right) \widehat{N}_{z_i, p_0}\left(dn_1\right) ds_1 ds_2$$

$$+ \int\limits_{s_1 < b_i^1} \pi_i^{(2)}\left(b_i^{2*}; v_i, z_i, p_0, s_1\right) f_{p_0}^1\left(s_1, s_2, n_1\right) \widehat{N}_{z_i, p_0}\left(dn_1\right) ds_1 ds_2.$$

I characterize the equilibrium in this game without deriving the exact functional forms for best responses and equilibrium actions. Note that in period 3 the payoff function of the bidder is strictly increasing in individual bid. Moreover, the bidding problem in the last period becomes static. Therefore, the equilibrium in the last period will coincide with the equilibrium in the static second-price auction where $b_i^3 = v_i$. To find the best response in the second and first periods one can write down the first-order condition. Specifically, in the second period:

$$\frac{\partial \pi_i^{(2)}\left(b_i^2; v_i, z_i, p_0, p_1\right)}{\partial b_i^2}$$

$$= \frac{\partial \pi_i^{(3)}\left(v_i; v_i, z_i, p_0, p_1, b_i^2\right)}{\partial p_2} \int\limits_{\substack{s_2 < b_i^2 \\ s_1 > b_i^2}} f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) \widehat{N}_{z_i, p_0, p_1}\left(dn_1, dn_2\right) ds_1 ds_2$$

$$-\pi_i^{(3)}\left(v_i; v_i, z_i, p_0, p_1, b_i^2\right) \int\limits_{s_2 > b_i^2} f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) \widehat{N}_{z_i, p_0, p_1}\left(dn_1, dn_2\right) ds_1 ds_2$$

$$+\pi_i^{(3)}\left(b_i^{3*}; v_i, z_i, p_0, p_1, b_i^2\right) \int\limits_{s_1 < b_i^2} f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) \widehat{N}_{z_i, p_0, p_1}\left(dn_1, dn_2\right) ds_1 ds_2.$$

Note that if $\frac{\partial \pi^3}{\partial p_2} < 0$, then the first-order condition will have an interior solution. To see this, note that if b_i^2 approaches zero, then the derivative of expected profit is negative, while if b_i^2 is very large, it becomes positive (given that the density of bid distribution has infinite support).

This setup can be represented in a more general form. Suppose that $\xi_t = (s_1^t,\ s_t^2)$ is a vector of the first and the second-order statistic for the distribution of bids in period t and

$$h\left(b, \xi_t\right) = \begin{cases} s_2^t - p_{t-2}, & \text{if } \ b < s_2^t, \\ b - p_{t-2}, & \text{if } \ s_2^t < b < s_1^t, \\ s_1^t - p_{t-2}, & \text{if } \ b > s_1^t. \end{cases}$$

Conditional distribution

$$f_\theta\left(\xi_t | p_{t-1}\right) = \begin{cases} \int f_{p_0}^1\left(s_1, s_2, n_1\right) N_\theta\left(dn_1\right), & \text{if } t = 1 \\ \int f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) N_\theta\left(dn_1, dn_2\right), & \text{if } t = 2 \\ \int f_{p_2}^3\left(s_1, s_2, n_1, n_2, n_3\right) N_\theta\left(dn_1, dn_2, dn_3\right), \end{cases} \tag{1.1}$$

where θ is known generates the true transition density for the players' entry. Conditional distribution

$$\hat{f}_{z_i}\left(\xi_t | p_{t-1}\right) = \begin{cases} \int f_{p_0}^1\left(s_1, s_2, n_1\right) \widehat{N}_{z_i, p_0}\left(dn_1\right), & \text{if } t = 1 \\ \int f_{p_1}^2\left(s_1, s_2, n_1, n_2\right) \widehat{N}_{z_i, p_0, p_1, p_2}\left(dn_1, dn_2\right), & \text{if } t = 2 \\ \int f_{p_2}^3\left(s_1, s_2, n_1, n_2, n_3\right) \widehat{N}_{z_i, p_0, p_1, p_2}\left(dn_1, dn_2, dn_3\right), \end{cases} \tag{1.2}$$

can be associated with bidder i's belief regarding the distribution of element ξ_t given bidder's type and current information. I assume that there exists a local sufficient statistic $\hat{\theta}$ such that

$$\hat{f}_{z_i}\left(\xi_t | p_{t-1}\right) = f_{\hat{\theta}}\left(\xi_t | p_{t-1}\right) + o\left(\left|\hat{\theta} - \theta\right|\right).$$

I will call $\hat{\theta}$ "belief regarding the visitbility". In this case the game can be represented in the following form. The auction starts at period $t = 1$ and ends at period $t = 3$. Players enter into the auction following Poisson process with frequency $\lambda(t, p_t, \theta)$. Poisson entry process generates the state variable ξ_t^i observed by each player i and has a density $f_\theta\left(\cdot|p_{t-1}\right)$. Actions of players are real positive bids, and the strategy of each player maps sequences of prices up to time t to bids submitted at time t. Player i is characterized by her valuation v_i and initial beliefs regarding the visibility of the auction z_i (player's type). In periods 1 to 3 she computes the sufficient statistic for the distribution of the state variable ξ_t, $\hat{\theta}_i$ given the path of prices and her type z_i. Player i forms beliefs regarding the distribution of the state variable ξ_t characterized by the density $\hat{f}_{\hat{\theta}_i}\left(\cdot|p_{t-1}\right)$. The price in period t for player i is

$$p_t = p_{t-1} + h\left(b_i^t, \xi_t^i\right).$$

In equilibrium the following conditions will be satisfied:

- Bidders maximize expected payoff at the end of the auction given their beliefs regarding ξ_t^i
- Price is equal to the second-highest bid in each period
- Beliefs of players regarding ξ_t^i are consistent with the true distribution of ξ_t^i, i.e. $\hat{f}_{\hat{\theta}_i=\theta}\left(\cdot|p_{t-1}\right) = f_\theta\left(\cdot|p_{t-1}\right)$

The information structure in this environment allows the bidders to form beliefs regarding the distribution of the state variable ξ_t (the highest and the second highest bids among participating bidders) and compute iteratively the sufficient statistic $\hat{\theta}$.

This allows the bidders to avoid computing the distribution of the number of entering bidders in each period. I will use this convenient property reducing the dimensionality of decision problems of individual bidders.

1.3.3 Continuous-Time Model

Motivating example in the previous subsection demonstrates that strategic behavior in the dynamic auction environment can be described by the optimization problem for an individual bidder who has consistent beliefs about the dynamics of the price. Here I will use the results from [San07] to describe this model in continuous time. The continuous-time auction model considered here describes a price process that starts at time 0 and ends at a time T. Consider the strategic behavior of a single bidder i who competes against her rivals by submitting multiple bids $b_i \in \mathbb{R}_+$ at any moment of time $t \in [0, T]$.

The auction has a second-price structure, so that the price of the object - $p \in \mathbb{R}_+$ - at a given time t is equal to the second highest bid among the bids submitted before time t. A bidder observes neither the identities nor the number of her rivals, and I denote the number of bidders at time t by N_t. She knows only that entry into the auction is characterized by a Poisson process with frequency $\lambda(t, p, \theta_0)$, a function of time, price and the visibility parameter θ_0[13], while its functional form is a common knowledge. This function is the instantaneous demand described above, whose dependence on price is motivated by equilibrium search considerations.

The visibility parameter θ_0 is exogenous and fixed for a specific auction. The set of possible visibility values Θ is convex, compact, and known by the bidders. Although θ_0 is not observed, bidders have initial beliefs regarding its value. These

[13] [BFS03] observed that the number of hits at popular Internet sites can be described by a Poisson process

beliefs take the form of normal priors, truncated to the set Θ, with means μ_θ and variances σ_θ^2 drawn from the distributions $G_\mu(\cdot)$ and $G_\sigma(\cdot)$ respectively. Bidders with small mean-squared errors of the initial belief $(\mu_\theta - \theta_0)^2 + \sigma_\theta^2$ are considered "more informed", while bidders with large mean-squared errors are "less informed".

In addition to being uncertain about the visibility of the auction the bidders also have only imperfect observations of the price process. Due to the inability to continuously monitor the price in the auction (late at night, for example), a price change occurring at time t will be observed by the bidder at time $t + \epsilon$ where ϵ is independent and drawn from the same (non-negative) distribution across the bidders. Bidders are assumed to observe their own ϵ and it is constant throughout the auction for the given bidder. The instantaneous demand for an individual bidder with an observation delay ϵ will be denoted $\lambda_\epsilon(t,\, p,\, \theta)$. For example, one of possible forms for the transformation of $\lambda(\cdot)$ into $\lambda_\epsilon(\cdot)$ is $\lambda_\epsilon(t, p, \theta) = \alpha e^{a\epsilon} \lambda(t, p, \theta)$ for fixed a and α.

The bidder's valuation for the object is v and she is risk-neutral: hence if she obtains the object for price p, then her utility is $v - p$. The valuations of the bidders are independently drawn from the distribution $F(v)$. The bidder maximizes the expected utility[14] from winning the auction $E_0 \{(v - p_T)\, \mathbf{1}[b_T > p_T]\}$. This expected utility is positive only if the bid of the bidder under consideration at the end of the auction is the highest (and, thus, is higher than the price of the object equal to the second highest bid). The strategy of the representative bidder is characterized by a bidding function $b_{v\,\epsilon}(t,\, p,\, \mu,\, \sigma)$, which gives the optimal bid value at time t for a

[14]It can be argued that in some cases individuals might bid on eBay for reasons other than maximizing the expected surplus from winning when, for instance, they like gambling. A variety of possible other targets are discussed in [Kag95]. In this paper, I focus only on maximization of expected surplus.

bidder with valuation v and observation delay ϵ if the price of the object is p and the beliefs of the bidder about the visibility have mean μ and variance σ. To facilitate derivation, I suggest the following decomposition of the bidding function:

$$b_{v\epsilon}(t, p, \mu, \sigma) = p + \eta(t, p, \mu, \sigma). \tag{1.3}$$

The function $\eta(\cdot)$ will be referred to as the bid increment. In general $\eta(\cdot)$ depends on ϵ and v, but to facilitate the notation we will dropping these indices wherever it clear which bidder is described. As noted above, the bidder wins the auction if her bid increment is positive at the end of the auction.

Definition 1.1. *A pure strategy of the bidder i η_i is a stochastic process with sample paths in $\mathbb{R}_+$ which is progressively measurable with respect to the filter generated by the price process.*

The price is always equal to the second highest bid. From the perspective of a particular bidder who submits at least a minimum required bid higher than the current price, the price is equal to the minimum of her own bid and the highest bid of her rivals. In this way, if she submits a bid and is the current highest bidder, then the current price is determined by the highest bid of the remaining bidders. A new bidder determines the price if her bid is between the current highest bid and the second highest bid. Any particular bid of a bidder can determine the price only once per auction.

Bidders form beliefs regarding the distributions of the price process at each instant. If bidder i has full information regarding parameter θ_0 in the entry process, her beliefs of bidder i about the price movement can be characterized by the finite-dimensional distributions of the stochastic process:

$$dp_{t,\epsilon_i} = h\left(t,\, p_t,\, \eta_i\right)\, dJ_{\epsilon_i}\left(t,\, p_t,\, \theta_0\right), \tag{1.4}$$

where dp_t are changes in the price over infinitesimal time intervals dt, and $dJ_\epsilon\left(t,\, p_t,\, \theta_0\right)$ are increments of the Poisson process with frequency $\lambda_\epsilon\left(t,\, p_t,\, \theta_0\right)$.[15] This equation reveals that the price, as observed by an individual bidder, evolves in jumps. The size of the jump at time t is equal to $h\left(t,\, p_t,\, \eta_i\right)$ where the bid increment η is defined at (1.3) while the timing of the jumps is governed by the Poisson process $J_\epsilon(\cdot)$, so that $n_t^J = \int\limits_0^t dJ_\epsilon\left(\tau,\, p_\tau,\, \theta_0\right)$ is the number of price jumps from the beginning of the auction up to time t[16]. Thus if n_t^J is the number of price jumps up to t and t_i are the times when jumps occur for $i = 1, \ldots, n_t^J$, then the price can equivalently be written as:

$$p_{t,\epsilon_i} = \sum_{i=1}^{n_t^J} h\left(t_i,\, p_{t_i},\, \eta_i\right). \tag{1.5}$$

This expression shows that the price of the object at time t is equal to the total sum of the price jumps up to time t.

Second, the timing of the price jumps is determined by the rest of the bidders,

[15] A more formal treatment of the conditions providing the existence of the function in the described model is given in Chapter 3. I will be using footnotes in this section to give short comments regarding the formal properties of the functions in the model. To describe the model formally I consider a process defined on a complete probability space $(\Omega,\, \mathbb{F},\, \mathbf{Q})$ and $\{\mathfrak{F}_t,\, t \in [0,\, T]\}$, a filter such that $\mathfrak{F}_t \subset \mathbb{F}$ and $J(t,\, x)$, and a Poisson measure for $t \in [0,\, T]$ adapted to filter $\{\mathfrak{F}_t\}$.

[16] Formally we need to make sure that the solution to this stochastic differential equation exists for given $h(\cdot)$ in terms of a stochastic Itô integral. For that purpose we impose two requirements on $h(\cdot)$: (i) $\int_0^T \left|h(t,\, x_t, \eta)\right|^2\, \lambda_\epsilon(t,\, x_t)\, dt < +\infty$ with probability 1; (ii) The compensated process $\int\limits_0^t h(t,\, x_t, \eta)\, \overline{J_\epsilon(dt,\, dx)}$ is a local square integrable martingale adapted to the filter $\{\mathfrak{F}_t\}$ with piecewise - continuous sample functions.

while the observed timing is contaminated by observation delay error. In the absence of observation delays, bidders would coordinate their bidding such that all participants would bid at the same time when new information arrived.

Equation (1.4) contains the visibility parameter θ_0, which is unobserved by the bidders. Because a bidder forms her bidding strategy to optimally influence the auction price, it is important that she predict the visibility parameter precisely. This prediction incorporates prior information, which the bidders can have from previous bidding experience and from observing the behavior of the price in the auction. In this way, as the price evolves during the auction, the precision of the bidder's estimate of the visibility of the auction increases. This process can be the result of strategic learning. One way to describe the estimation of the visibility parameter by the bidder is to consider Bayesian updating of the prior beliefs given the stochastic movement of the price.[17]

To compute the dynamics of these beliefs I use the linear filtration method developed in [VS77]. While an optimal linear filter (as shown in [LS01]) is infinite-dimensional, for convenience of computing the bidder's estimate of the visibility of the auction, I restrict the analysis to the first two moments[18]. Chapter 3 shows that

[17] The specific problems of construction of Bayesian estimates for Poisson - type processes are described, for example in [Kut98] and [Kar86]. For the Gaussian processes the solution to nonlinear mean-square inference problem may be found, for example in [LS01] in the form of the system of stochastic differential equations. In a more general case of point-stochastic processes, a solution becomes complicated, relies heavily on martingale properties of the underlying stochastic processes and, more importantly, becomes a computationally intensive problem. A constructive way of building linear Bayesian forecasts for the parameters of a Poisson process with a variable frequency is shown in [Gra72]. The generalization of Grandell's method follows from the integral representation of martingales and is discussed in detail, for example, in [LS01]. Here I will discuss an easier and more intuitive approach which, despite being less rigorous, provides a straightforward way to construct linear Bayesian estimates for the unknown parameters.

[18] In other words if $\mathfrak{F}_{xt}$ are the σ-algebras generated by the sample trajectories of the price process, then my assumption is that $E\{(\theta_t - \theta^*)^3 | \mathfrak{F}_{xt}\} = \sigma_3$ for any $t \in [0, T]$.

the posterior distribution for θ_0 can be described by a normal posterior distribution with mean μ_t, and variance parameter σ_t. The mean and variance parameter of the bidder's belief regarding the auction's visibility change over time according to the following system of stochastic differential equations:

$$\begin{aligned} d\mu_t &= \tfrac{\partial\lambda(t,p_t,\mu_t)}{\partial\theta}\tfrac{\sigma_t}{\psi(t,p_t,\mu_t)}\left\{dp_t - \psi(t,\,p_t,\,\mu_t)\,dt\right\}, \\ d\sigma_t &= -\tfrac{1}{\lambda(t,p_t,\mu_t)}\left(\tfrac{\partial\lambda(t,p_t,\mu_t)}{\partial\theta}\right)^2\sigma_t^2\,dt. \end{aligned} \tag{1.6}$$

In this expression $\psi\left(t,\,p_t,\,\mu_t\right) = h\left(t,\,p_t,\,\eta_t\left(t,p_t,\mu_t,\sigma_t\right)\right)\lambda_\epsilon\left(t,\,p_t,\,\mu_t\right)$ characterizes the expected growth rate of price per unit of time.

1.3.4 Equilibrium

In reality, auction markets are complex and involve repeated strategic interactions between buyers, sellers and the auction company. The equilibrium that I examine in this paper considers the behavior of bidders in a single auction and takes as given the behavior of the sellers and the auction engine itself.

In my model, the bidder forms beliefs regarding the distribution of the price at each instant. In equilibrium, these beliefs should be consistent with the true price distribution generated by aggregate behavior of multiple bidders, while bidders have to maximize their expected payoffs from winning the auction. To formalize the notion of the equilibrium, I use the results in [San07] and combine them with the Bayesian structure of the game that I consider in this paper. Following [San07] I call the equilibrium here a perfect public equilibrium. The components of equilibrium are: (i) entry of new bidders, (i) types of new bidders and their beliefs regarding the price process, (iii) a profile of strategies of bidders who have entered into the auction up to time t.

Definition 1.2. *A profile of strategies* $\left\{(\eta_{i,\tau})_{i=1}^{N_\tau}\right\}_{\tau=0}^{T}$ *along with beliefs of bidders regarding the price process* $\mathbf{P}\left\{p_t < x,\ J\left(t, p_t, \mu_t^i\right) < y\right\}$ *and means and variances of the beliefs of the bidders regarding the visitbility* $\{\mu_t^i,\ \sigma_t^i\}_{i=1}^{N_t}$ *constitute a perfect public equilibrium if for each* $t \in [0,\, T]$*:*

- *The entry of the bidders is a Poisson process with the frequency* $\lambda(t,\, p,\, \theta_0)$ *for a given* $\theta_0 \in \Theta$ *so that the number of bidders who entered up to time* t *is* N_t *and is unobserved by the participating bidders*

- *Observable public histories consist of price paths up to time* t

- *Bidders' types include valuations, initial beliefs regarding the visitbility, and observation delays. The valuations of the bidders are drawn from the distribution* $v^i \sim F(\cdot)$, *the initial parameters of the beliefs are* $\mu_0^i \sim G_\mu(\cdot)$ *and* $\sigma_0^i \sim G_\sigma(\cdot)$, *and the observation delay errors are uniformly distributed* $\epsilon^i \sim U[0,\, 1]$ *for all bidders* $i = 1, \ldots, N_t$ *and for each* $t \in [0,\, T]$[19]

- *Each bidder* i *maximizes the expected payoff from winning the auction given her beliefs regarding the finite-dimensional distributions of the price process*

$$\mathbf{P}\left\{\int h\left(t,\, p_t,\, \eta_i\right)\, dJ_{\epsilon_i}\left(t,\, p_t,\, \mu_t^i\right) < x,\ J\left(t, p_t, \mu_t^i\right) < y\right\}$$

  and her type

- *A profile of actions* $\left\{(\eta_{i,\tau})_{i=1}^{N_\tau}\right\}_{\tau=0}^{T}$ *along with the Poisson entry process* N_t *generate the price process such that its finite-dimensional distributions are consis-*

[19]As I will show in the next section, the distribution of the observation error and the frequency of entry are not separately identified. For this reason, I assume here that the distribution of observation errors is uniform to estimate the frequency of entry.

tent with beliefs of bidders regarding the price process:

$$\mathbf{P}\Big\{\int_0^t \max_{i=1,\dots,N_t,\ i\neq J} \eta_{i,t}<x, J(t,p_t,\theta_0)<y\Big\}$$
$$=\mathbf{P}\Big\{\int h(t,p_t,\eta_i)\,dJ_{\epsilon_i}\big(t,p_t,\mu_t^i\big)<x,\ J\big(t,p_t,\mu_t^i\big)<y\Big\}\Big|_{\substack{\mu_{i,t}=\theta_0\\ \epsilon_i=0}}. \tag{1.7}$$

for $J = argmax_j \eta_{j,t}$.

It is important to note that in my model bidders can submit multiple bids, and equation (1.7) refers to the last bid of every bidder who entered into the auction up to time t. This equilibrium concept suggests that the price increase at time t is driven by the second order statistic of the bid increment function computed at price p_t. In this way, a competitive dynamic equilibrium is a solution of the stochastic differential equation (1.7).

Let me first consider the steps that establish the existence of equilibrium in the continuous-time auction model. It is straightforward to establish existence and uniqueness of the fixed point by considering the individual optimization problem of the bidder. An equilibrium in this setting corresponds to the solution of a collection of individual bidding problems of the participating bidders. This collection is a system of partial differential equations describing the law of motion of the value functions of the bidders in the auction. Chapter 3 shows that, by certain transformations, this system of differential equations can be represented as a contraction mapping. As a result, existence of the equilibrium is straightforward.

To prove uniqueness of the equilibrium, we must account for entry of bidders into the auction. By assumption, this entry is driven by a Poisson process with variable frequency. If this frequency is finite for any value of the price, time, and visibility parameter, then the number of bidders entering into the auction per unit of time is finite with probability one. This means that the equilibrium in the continuous-time auction can be represented as a finite collection of optimal bidding problems of individual bidders with probability one. As a result, this collection has a unique fixed point that is the equilibrium in the continuous-time auction. A more detailed treatment of this result is presented in Chapter 3. I will use this uniqueness result to describe the likelihood of the continuous-time auction model.

1.3.5 Individual Bidding Behavior with Consistent Beliefs

In the previous subsection, I characterized the public perfect equilibrium with consistent beliefs which generates the dynamic behavior in the considered environment. I assumed that bidders maximize their expected utilities from winning the auction by submitting optimal bids. In equilibrium beliefs of individual bidders regarding the price distribution should be consistent with the true distribution of the price. In the following two subsections I will describe individual best responses. In this subsection I will provide a method for finding individual best responses of the bidders provided that they have consistent beliefs regarding the price distribution. In case of consistent beliefs, bidding strategies are described by the bid increment functions $\eta\left(t, p, \mu, \sigma\right) \in \mathbb{R}_+$ defined as the difference between the submitted bid of the bidder and the current price. Expected utility maximization is constrained by the dynamics of the price, described by the stochastic differential equation (1.4). The visibility of the auction θ_0 is unobserved by the bidders, but they try to infer its value from observing the actual movement of price in the auction. I assume that the bidders infer the visibility of the auction by forming priors and update their beliefs as the price in the auction changes. This allows me to write the problem of the bidder as:

$$\begin{aligned}
&\max_{\eta(\cdot)}\ E_0\{(v-p_T)\,\mathbf{1}\,[\eta_T>0]\}\\
&dp_{t,\epsilon}=h\,(t,\,p_t,\,\eta)\ dJ_\epsilon(t,\,p_t,\,\mu_t),\\
&d\mu_t=\tfrac{\partial\lambda(t,p_t,\mu_t)}{\partial\theta}\tfrac{\sigma_t}{\psi(t,p_t,\mu_t)}\left\{dp_t-\psi(t,\,p_t,\,\mu_t)\,dt\right\},\\
&d\sigma_t=-\tfrac{1}{\lambda(t,p_t,\mu_t)}\left(\tfrac{\partial\lambda(t,p_t,\mu_t)}{\partial\theta}\right)^2\sigma_t^2\,dt.\\
&\theta_t\big|_{t=0}\sim N\left(\mu_0,\,\sigma_0\right),\quad p_t\big|_{t=0}=p_0.
\end{aligned}\tag{1.8}$$

The first equation of this dynamic optimization problem represents the objective of the bidder, which is the expected utility of winning the auction. This objective reflects the fact that the bidder obtains a positive utility from the auction only if she wins it (so that her bid is the highest, implying that the bid increment at the end of the auction is strictly positive $\eta_T > 0$). The expectation $E_t\,[\cdot]$ is taken over the information set up to time t which includes the paths $\{p_t,\,\mu_t,\,\sigma_t\}_{t=0}^{T}$.

The second equation demonstrates the price process corresponding to beliefs of an individual bidder regarding the price movement. It describes the dynamics of the price as a jump process with Poisson-driven jumps. The frequency of price jumps is $\lambda_\epsilon\,(t,\,p_t,\,\mu_t)$. The bidders observe the price jumps with a random delay ϵ. The mean of the belief of the bidder about the visibility of the auction is μ_t.

The third and the fourth equations represent the evolution of the mean and variance of the bidder's beliefs regarding visibility. These beliefs are driven by the price changes so that the mean of the distribution shifts when the auction price jumps. Moreover, if the expected price growth rate is elastic with respect to visibility, then the variance of the bidder's beliefs will decrease over time.

The last line in the dynamic optimization problem (3.6) sets the initial conditions. Bidders' initial beliefs about the visibility parameter , θ_0, take the form of a normal distribution with mean μ_0 and variance σ_0. The price at the beginning of the auction is the reserve price for the object set by the seller. It is assumed to be observable by the bidders and equal to p_0.

1.3.6 Individual Best Responses with Consistent Beliefs

My approach to solving the problem (3.6) uses the Bellman equation formulation, which has been studied extensively in the literature on stochastic dynamic optimal control. Here, we can define the value function of the bidder as:

$$V(t, p, \mu, \sigma) = E_t \left\{ (v - p_T) \mathbf{1} [\eta_T > 0] \right\}.$$

The function $V(\cdot)$ specifies the expected surplus of the bidder from winning the auction given that, at time $t \in [0, T]$, the price in the auction is p and the belief about the visibility of the auction has mean μ and variance σ. The value function at the end of the auction is the utility from winning the auction (to win the auction the bidder must have the highest bid, but the price is determined by the second highest bid, so $\eta_T = b - p_T > 0$). The vector of state variables $(p_t, \mu_t, \sigma_t)'$ forms a sufficient statistic for the dynamics information embedded in the entire price path up to time t. The dynamic evolution for the state variables $(p_t, \mu_t, \sigma_t)'$ over time is a Markov process.

The optimal behavior of the representative bidder is derived by applying the Itô calculus to the value function. Intuitively, the derivation uses the fact (due to the Bellman principle) that the expected surplus of the bidder at time t is equal to the maximum over all possible bid increments of the expected surplus at time $t + dt$. The expectation is taken over the distribution of all possible price changes in the interval of time dt. The bidders optimally influence the price process to maximize their expected payoff. As a result, we can represent the law of motion of the value function of the bidder as:

$$\begin{aligned}
&\frac{\partial V(t,p,\mu,\sigma)}{\partial t} + \sup_{\eta\in\Xi}\Bigg[-\frac{\sigma}{h}\frac{\partial\lambda}{\partial\theta}\frac{\partial V}{\partial\mu} - \frac{\sigma^2}{\lambda}\left(\frac{\partial\lambda}{\partial\theta}\right)^2\frac{\partial V}{\partial\sigma} + \\
&+V\left(t,\, p+h\left(t,p,\eta\right),\, \mu+\frac{\partial\lambda}{\partial\theta}\frac{\sigma}{\lambda},\, \sigma\right)\lambda_\epsilon\left(p+h\left(t,p,\eta\right),\, \mu+\frac{\partial\lambda}{\partial\theta}\frac{\sigma}{\lambda},\, t\right) - \\
&-V(t,\, p,\, \mu,\, \sigma)\,\lambda_\epsilon\left(p,\, \mu,\, t\right)\Bigg] = 0,
\end{aligned} \tag{1.9}$$

$$V(T,\, p,\, \mu,\, \sigma) = \sup_{\eta\in\Xi}\left\{(v-p)\,\mathbf{1}\left[\eta>0\right]\right\}. \tag{1.10}$$

In this equation, the space of the control functions Ξ limits my analysis to the bounded and piecewise-continuous bid increment functions, simplifying further derivations[20]. Equation (3.7) is a partial differential equation for the value function of the bidder $V(\cdot)$ with boundary condition (1.10) which implies that the value function of the bidder at the end of the auction has to be equal to her utility from winning the auctioned object[21]. Note that the boundary condition in (3.7) implies that the optimal bid increment at the end of the auction can take a range of values $\eta_{v\epsilon}\left(T,p,\mu,\sigma\right) \in (0,\, +\infty)$. If we analyze the optimal strategy in the interval $[T-\tau,\, T]$ then the first equation in the boundary problem (3.7) suggests that as $\tau \to 0$ then $v-p$ is optimal for $\eta_{v\epsilon}\left(T,p,\mu,\sigma\right)$ if $v>p$. Therefore, it is optimal for the bidder to submit a bid equal to her true valuation at the last moment of the auction, analogous to the behavior

[20] In fact I require that $\eta\in\Xi$ are continuous and differentiable almost everywhere functions defined on $[0,\, T]\times[0,\, \overline{X}]\times\Theta\times\Sigma$. I assume that these functions are bounded by some $\overline{X}<\infty$ and that $h(\cdot,\eta(\cdot))$ are measurable functions of finite variance with respect to the Poisson measure.

[21] Technically speaking, I have obtained a boundary problem for the expected surplus only. This determines the behavior of the function on one of the boundaries but does not provide the information about the derivatives. Such a boundary problem is often classified as a Dirichlet problem.

in a static second-price auction.

1.3.7 Discussion of the Individual Bidding Problem

A collection of individual bidding problems generates a dynamic public perfect equilibrium. The individual bidding problem for the bidder with consistent beliefs regarding the price process is represented by the boundary problem (3.7) and (1.10). I will first describe the general properties of this problem and then discuss the interpretation of the individual components of the individual decision problem.

The problem (3.7), (1.10) has a two-component structure. The component in the supremum describes optimal bidding at time t. This component allows me to compute the bidding strategy for a fixed instant as it produces the optimal response of the bidder at time t to the current price, and the mean and variance of the bidder's beliefs regarding auction visibility. After the optimal bidding strategy is computed at time t, I can compute the time derivative of the value function of the bidder and, as a result, obtain the value function for the previous instant of time[22]. Having discussed the general structure of the bidder's problem, let us now proceed with the interpretation of the individual components describing bidding behavior at a specific

[22]This two component structure of the problem naturally enables me to use the standard Euler algorithm to compute the optimal behavior of the bidder. The algorithm would begin from the boundary condition which states that at the end of the auction (T) the value function of the bidder is equal to her utility from winning the auction. Then I make a step backwards in time τ and recompute the value function at time $T - \tau$. The step size τ determines the precision of the computed value function. For this value function I compute the optimal bid increment of the bidder and evaluate the time derivative of her value function. This process repeats until it reaches the beginning time 0. This approach has a significant computational advantage over the value function iteration approach which is used to compute the optimal behavior in an infinite horizon discrete time optimization problem. As compared to the value function iterations, there is no need to prove the convergence of the iteration method and wait until the convergence is reached. In my case the number of steps is fixed and is equal to the number of grid points of time while the convergence of Euler's steps for the differential equations is well known. Moreover, such an approach allows me to use a continuous state space which is not always feasible in the discrete time case.

point of time in the auction.

The last two terms in equation (3.7) describe the main strategic tradeoffs faced by each bidder. These strategic tradeoffs highlight the role of visibility in the bidder's behavior: higher visibility will in general imply that the bidders will bid more aggressively. The information about the visibility influences the bidding frequency and allows me to distinguish the more informed bidders from the less informed ones. Let me consider these two terms by components. The component

$$V\left(t, p+h\left(t, p, \eta\right), \mu, \sigma\right) \lambda_{\epsilon}\left(t, p+h\left(t, p, \eta\right), \mu\right) - V(t, p, \mu, \sigma)\lambda_{\epsilon}\left(t, p, \mu\right)$$

describes the expected change in the value function due to the price jump. The tradeoff of the bidder is between bidding early and bidding late. Bidding early deters entry, decreasing the probability of price jumps and resulting in a higher value function at the end of the auction. On the other hand, the direct effect of bidding early is to raise the price throughout the auction, which decreases the value function of the bidder at the end of the auction. The optimal time to bid occurs when these two effects offset each other. The movement of the value function is determined by the product of the current value function and the rate of entry into the auction $\lambda_{\epsilon}(\cdot)$ (instantaneous demand). As a result, the bidder maximizes the total value of the entry rate (value times demand) and extracts a surplus from bidding by choosing the optimal point on the demand schedule by means of bidding early. This behavior constitutes entry deterrence.

The component

$$V\left(t, p, \mu+\frac{\partial \lambda}{\partial \theta}\frac{\sigma}{\psi}, \sigma\right) \lambda_{\epsilon}\left(t, p, \mu+\frac{\partial \lambda}{\partial \theta}\frac{\sigma}{\psi}\right) - V(t, p, \mu, \sigma)\, \lambda_{\epsilon}\left(t, p, \mu\right)$$

reveals the role of information in the strategic decision of the bidders. Note that the size of this component depends on the variance of a bidder's belief regarding the visibility of the auction, so that the effect will be different for bidders with different variances of prior beliefs. If the variance of the prior is large, the bidder does not have much information about the visibility of a specific auction. As a result, the value function of the bidder will significantly depend on the behavior of price. As the variance of the bidder's belief diminishes (towards the end of the auction), the value function will become less sensitive to the new information and the strategy of this bidder will converge to the strategy of the fully informed bidder. Therefore bidders with large prior variances will prefer rapid changes in price because this allows them to reduce the variance of their beliefs faster and extract a higher surplus from bidding. The other type of bidders includes those who initially have low variances of beliefs. In the limit, if the variance of the prior is zero, these bidders will not change their strategies as price changes. Specifically, if this type of bidder deters the entry of the other potential bidders at the beginning of the auction (or the instantaneous demand is initially low), she will wait until the end of the auction to submit her final bid since there are no incentives for her to bid more frequently. In this way, she will not allow her information about the visibility of the auction to be absorbed by the price. This type of behavior constitutes learning prevention.

The component $-\left\{\frac{\sigma}{\lambda}\frac{\partial\lambda}{\partial\theta}\right\}\frac{\partial V(t,p,\mu,\sigma)}{\partial\mu}$ can be interpreted as the effect of the mean of the bidder's beliefs on the value function. Specifically, this effect depends on the variance of beliefs and the sensitivity of the entry rate $\lambda(\cdot)$ with respect to the changes in the visibility. If the variance of the belief distribution σ is small, then the effect of a change in the mean of bidder's beliefs is small. Similarly, if the instantaneous demand λ is very sensitive to changes in visibility, then the value function of the bidder will be more sensitive to changes in the mean of bidder's beliefs.

The component $-\frac{\sigma^2}{\lambda}\left(\frac{\partial\lambda}{\partial\theta}\right)^2 \frac{\partial V(t,p,\mu,\sigma)}{\partial\sigma}$ can be interpreted the effect of the variance of a bidder's beliefs on her value function. This indicates that as bidders obtain more information about entry into the auction over time, their beliefs regarding visibility become closer to the truth. As a result, an increase in the variance of the bidder's belief will be followed by a decrease in the bidder's value function. Intuitively, this implies that the bidder values more precise information from the auction.

It is important to mention that my model is designed to analyze a partial equilibrium in a single auction within a context of a large auction market. In general equilibrium in the market visibility is endogenous characteristic: entry rate into particular auctions will depend on the total number of bidders in the market, the number of sellers, and characteristics of an auction (such as a position in the search contents, spelling errors in the name, number of cross-listed items).

1.3.8 Identification

The continuous-time auction model considered in this paper has a complex structure. In the theoretical section, I represented this structure in an algorithmic form which allowed me to describe the behavior of bidders as a reaction to a stochastically changing price. The reaction of the bidders is summarized by the second highest bid which is equal to the equilibrium price. In order to fit the model to the data, one makes a "guess" about the structural functions of the model (the instantaneous demand, the size of the price jumps, the distribution of valuations, and the distribution of bidders' beliefs), simulates the optimal behavior of the bidders, computes the second highest bid at each instant, and matches the path of the simulated second highest bid to the price path that is actually observed. In this section I verify that the model is identified and the result of the specified matching procedure establishes a one-to-one correspondence between the simulated distribution and structural parameters.

Definition 1.3. *The continuous-time auction model of the continuous-time auction is identified if there is a one-to-one correspondence between the joint probability distribution of price jump sizes and the jump times*
$\mathbf{P}\left\{\int h\left(t,p_t,\eta\right)\,dJ\left(t,p_t,\theta\right)<\pi,\ \int dJ\left(t,p,\theta\right)<\tau\right\}$ *and a specific collection of characteristics of the theoretical model: instantaneous demand function* $\lambda\left(t,\,p_t,\,\theta\right)$*, price jump size function* $h\left(t,\,p_t,\,\eta\right)$*, the distribution of valuations* $F\left(v\right)$*, and the distributions of initial beliefs* $G_\mu\left(\mu_0\right)$ *and* $G_\sigma\left(\sigma_0\right)$

This definition implies that if one simulates the equilibrium behavior of the bidders, this will also determine the characteristics of the theoretical model that produce the same simulated distribution of prices. According to this definition, identification is achieved if no other set of characteristics of the model produces the same distribution of simulated prices.

My strategy now will be to provide a set of conditions that allow me to demonstrate that the model is non-parametrically identified using Definition 3.1. As a result, the potential outcome of the estimation procedure will be the complete functional form of the instantaneous demand function, the price jump function, and the distributions of bidders' valuations and beliefs. From the point of view of Definition 3.1, the model will be identified if two conditions hold. First, the observable distributions should contain sufficient information about the structural functions of the model. The known observational outcomes are the joint distribution of timing and sizes of the price jumps. The structural functions of the model can be recovered by using some method of inversion of the joint distribution of timing and sizes of the price jumps. As a result, identification of the model is partially assured by the structure of the data and partially assured by the functional form assumptions that allow me to invert the estimated distribution.

The data collected from eBay contain the complete price paths for multiple auctions. My random data collection methodology assures that there is no selection bias in the distribution of the visibility in the sample of auctions that I am using for estimation. Namely, the collected sample of auctions will be considered free of selection bias (i.e. I did not pre-select the auctions with certain visibility). Moreover, I will assume that the data are collected for very uniform and similar items, guaranteeing that the functions driving the continuous-time auction (the instantaneous demand, the size of the price jumps, and the distribution functions) are the same across auctions, conditional on individual-specific covariates. These data collection assumptions assure that the dataset possesses a certain degree of independence and homogeneity, so that observing repetitions of the auction with certain parameters is equivalent to observing a cross-section of simultaneous similar auctions. As a result, my data have two dimensions: on the one hand, I have a complete record of individual bids corresponding to the continuous-time observations of the price process over time. On the other hand, for each instant I can see the distribution of prices across the auctions. In this way, under certain assumptions I will be able to disentangle the characteristics which are constant over time by looking at the cross sectional dimension, and then recover the remaining characteristics from the data over time.

I will now formulate a set of assumptions which will allow me to establish identification of the theoretical model.

- *Assumption 1.*

The support of the joint distribution of prices and the price jumps

$$\mathbf{P}\left\{\int_0^t h\left(t,p_t,\eta\right)\,dJ\left(t,p_t,\theta\right)<\pi,\ \int_0^t dJ\left(t,p,\theta\right)<\tau\right\}$$

is a convex compact set for any $t\in[0,\,T]$.

- *Assumption 2.*

  The instantaneous demand function observed by the bidders $\lambda_\epsilon\left(t,p,\theta\right)$ is decreasing with respect to the observation delay ϵ.

- *Assumption 3.*

  The instantaneous demand function observed by the bidder $\lambda_\epsilon\left(t,p,\theta\right)$ is strictly increasing in the visibility of the auction.

- *Assumption 4.*

  The bid increment function $\eta\left(t,p_t,v,\mu,\sigma\right)$ is non-decreasing in the bidder's valuation v and the mean of the prior belief of the bidder μ about the visibility of the auction, while bidder beliefs are independent from their valuations. Moreover, for each v there is a price $\overline{p}$ such that for all prices p higher than $\overline{p}$ we have $\eta\left(t,p,v,\mu,\sigma\right)=0$.

- *Assumption 5.*

  The bid increment function $\eta\left(t,p,v,\mu,\sigma\right)$ is less sensitive to changes in the price for bidders with smaller initial variance of beliefs about the visibility of the auction. That is for $\sigma'<\sigma$ and a sufficiently small $\Delta p>0$:

$$\left|\eta\left(t,p+\Delta p,v,\mu,\sigma'\right)-\eta\left(t,p,v,\mu,\sigma'\right)\right|\leq\left|\eta\left(t,p+\Delta p,v,\mu,\sigma\right)-\eta\left(t,p,v,\mu,\sigma\right)\right|.$$

Before beginning the proof of identification, I will discuss the role of each assumption. The first assumption suggests that there are no "holes" in the joint distribution of the number of price jumps and the price. In this way, if this distribution is observed at any moment of time, one can invert it to get the expected number of price jumps and the expected price.

The second assumption allows me to specify the direction in which the observation delay influences the instantaneous demand perceived by the individual bidders. This assumption suggests that the bidders with higher observation delays should be observing the price jumps less frequently.

The third assumption formalizes the notion of the visibility of the auction as a measure of "attractiveness" of a specific auction to the bidders. This enables me to estimate the visibility parameter from multiple auctions based on ranking them by the number of active bidders.

The fourth assumption allows me to distinguish aggressive bidding by bidders with high valuations from aggressive bidding by bidders with high beliefs regarding the visibility. This separation allows me to use information from different parts of the support of the bid increment distribution to estimate the distributions of valuations and the initial beliefs.

The fifth assumption allows me to identify bidder types based on their different reactions to the price jumps. Type indicates the precision of the initial information that the bidder has regarding the visibility of the auction.

The assumptions allow me to uniquely recover from the data the set of structural functions of my model. I will now outline the identification strategy and then proceed to a formal proof of identification.

For any given instantaneous demand function and the price jump size function, we can compute the optimal bidding function for individual bidders. The bidding function reflects the optimal bid value for the bidder given her valuation, beliefs, and the current price of the object for each moment of the auction. Existing papers such as [GPV00], [CGPV03], and [AH02] establish non-parametric identification of static structural auction models. My strategy in this paper extends the identification results to the continuous-time auction model. If one assumes that the bidding function is monotone with respect to the valuation at each moment of time and that the beliefs are independent from valuations, then the distribution of the number of active bidders across auctions given the price will reflect the distribution of valuations. As a result, a set of cross-sectional observations will identify the distribution of valuations.

Once the distribution of valuations is available, it becomes possible to identify the distribution of beliefs. Two facts are important for the identification of this distribution. First, according to Assumption 5 strategic learning occurs faster for bidders whose beliefs are close to the true visibility. We will then be able to identify the mean beliefs of the bidders by the distance between their observed bidding patterns given their information set and the optimal bidding pattern (computed for an auction with given visibility and given structural functions). Second, as the model predicts that bidders with more diffuse priors bid more frequently, we will be able to sort the bidders according to the relative number of their bids and, in this way, separate bidders by size of variance of initial beliefs. This allows me to identify the distribution of the bidders' beliefs from the observations across time by measuring the relative frequencies of bidding of different bidders and the distance of their bidding patterns from the optimal one.

Finally, given the distributions of valuations and beliefs, for a given instantaneous demand and price jump size function, I will be able to simulate the path of the second highest bid. If we match the distribution of the simulated second highest bid and the distribution of actually observed prices over time, we reduce the estimation of the model to the estimation of parameters of the compound Poisson process. Such estimation is described in detail in the literature (for instance, in [MW04], [LS01], and [Kar86]) and provides a unique outcome under conditions which are assumed to be satisfied here.

In addition to assumptions 1-5, I will assume that the distribution of the observation delay ϵ is known (and assume that it is uniform on $[0,1]$). I need to impose this assumption because this distribution is not separately identified from the instantaneous demand. Consider the individually perceived instantaneous demand $\lambda_\epsilon(t,\, p_t, \theta)$. Cumulative bidding behavior is associated with the "effective" frequency of price jumps $\int \lambda_\epsilon(t,\, p_t,\, \theta)\; dF(\epsilon)$ because ϵ was assumed to be independent across the auctions. For any differentiable monotone transformation $\varphi(\cdot)$ we can find a constant C and a function $\phi(\cdot)$ such that:

$$\int \lambda_\epsilon(t,\, p_t,\, \theta)\; dF(\epsilon) = \int \varphi\left(\lambda_\epsilon(t,\, p_t,\, \theta)\right) \frac{1}{C}\phi(\epsilon)\; dF(\epsilon),$$

and $\frac{1}{C}\int \phi(\epsilon)\; dF(\epsilon) = 1$. Thus, the distribution of the observation delay errors cannot be identified separately from the instantaneous demand function. For this reason I set this distribution to be uniform and impose the normalizations $\lambda_\epsilon(t,p,\theta)\big|_{\epsilon=0} = \lambda(t,p,\theta)$ and $\lambda_{\epsilon'}(t,p,\theta) < \lambda_\epsilon(t,p,\theta)$ if $\epsilon' > \epsilon$. This normalization implies that the instantaneous demand observed without any delay is equal to the actual instantaneous demand.

The formal identification argument is provided in the following theorem.

Theorem 1.1. *Suppose that we observe distributions of the form*

$$\mathbf{P}\left\{p_t,\, J_t\right\}, \quad \textit{for each } t \in [0,\, T]\,,$$

where p_t is the price in the auction at time t and J_t is the number of price jumps from the beginning to time t. Then under assumptions 1-5 and given that $\epsilon \sim U\,[0,\, 1]$ the following results are valid for identification from this collection of distributions:

- *The price jump size and the instantaneous demand function are identified;*
- *The distribution of valuations of the bidders is identified;*
- *The distributions of the beliefs of the bidders are identified up to the scale provided by the visibility of the auction: $G_\mu\left(\frac{\mu}{\theta_0}\right)$ and $G_\sigma\left(\frac{\sigma}{\theta_0}\right)$.*

Proof: The argument for the first statement follows from the generic identifiability of time-heterogeneous Poisson processes.

Now consider the last two statements. By my assumption, the distribution of the observation delay errors is uniform on $[0,\, 1]$. The instantaneous demand is assumed to be monotone with respect to the observation delay errors. In this case the conditional probability of the second highest bid given valuations and beliefs of bidders will be a monotone transformation of the bid increment function. For example, if the instantaneous demand is a multiplicative function of the observation delay:

$$\begin{aligned}
&\mathbf{P}_\epsilon\left(\max\nolimits_{(2)} \eta_{v\epsilon}\left(t, p, \mu_0, \sigma_0\right) < \pi\right) \\
&= \sum_{k=2}^{\overline{N}} \binom{\overline{N}}{k} \left[\frac{\pi}{\eta_{v\;\epsilon=1}(\cdot) - \eta_{v\;\epsilon=0}(\cdot)}\right]^k \left[1 - \frac{\pi}{\eta_{v\;\epsilon=1}(\cdot) - \eta_{v\;\epsilon=0}(\cdot)}\right]^{\overline{N}-k}
\end{aligned}$$

Here $\max_{(2)}$ denotes the second highest bid increment. The bid increment function is non-decreasing in the valuation of the bidder. This formula shows that the probability $\mathbf{P}_\epsilon\left(\max_{(2)} \eta_{v\epsilon}\left(t, p, \mu_0, \sigma_0, \right) < \pi\right)$ is a monotone one-to one map from the bid increment function to the the conditional distribution of the second highest bid. As a result, I can conclude that the conditional probability of the bid increment function is increasing in the valuations of the bidders. Then for two subsets $\mathbb{V}_1$ and $\mathbb{V}_2$ of $\mathbb{R}_+$ such that $\mathbb{V}_1 \leq \mathbb{V}_2$, the probability distributions of the second highest bid given the initial belief are also ordered:

$$\mathbf{P}_\epsilon\left(\sum_{\tau<t} \max_{(2)} \eta_{v_1\epsilon}\left(t, p, \mu_0, \sigma_0\right) < \pi\right) \leq \mathbf{P}_\epsilon\left(\sum_{\tau<t} \max_{(2)} \eta_{v_2\epsilon}\left(t, p, \mu_0, \sigma_0\right) < \pi\right),$$

for all $v_1 \in \mathbb{V}_1$ and $v_2 \in \mathbb{V}_2$, and the summation $\sum\limits_{\tau<t}$ is over all discrete moments of price jumps up to time t. This means that we can partition the distribution of valuations $F(v)$ into very small regions. If the unconditional distributions (with respect to ϵ) of bid increments in two distinct small regions are the same, the density of valuations in the region with higher valuations is lower. In this way we will be able to uncover the distribution of valuations over the entire state space using such indifference conditions.

Now let me consider whether I can provide a similar ordering for the conditional distribution of the second highest bid with respect to the parameters of the initial bidder's beliefs. By assumption 4, a higher mean of the belief distribution implies that the bidder expects a higher demand for the object. As it follows from the model, the bid increment function is non-decreasing with the instantaneous demand, implying that the bid increment function is non-decreasing in the mean of the bidder's belief.

In this case, for two subsets of Θ such that $\mathbb{T}_1 \leq \mathbb{T}_2$ we should have that:

$$\mathbf{P}_{\epsilon}\left(\sum_{\tau<t} \max_{(2)} \eta_{v\epsilon}\left(\tau, p_{\tau}, \mu_1, \sigma_0\right) < \pi\right) \leq \mathbf{P}_{\epsilon}\left(\sum_{\tau<t} \max_{(2)} \eta_{v\epsilon}\left(\tau, p_{\tau}, \mu_2, \sigma_0\right) < \pi\right),$$

such that $\mu_1 \in \mathbb{T}_1$ and $\mu_2 \in \mathbb{T}_2$.

The last parameter of interest is the variance of the bidders' beliefs at the beginning of the auction. By Assumption 5, the bid increment of bidders with more diffuse priors is more sensitive to the price than that of bidders with less diffuse priors. By monotonicity, the distribution of the second highest bid is more concentrated if the variance of the initial belief is smaller. In this way, the more informed bidders (those with lower variance of beliefs regarding visibility) will be using bidding strategies that are close to the optimal one. As a result, we can write that for two sets $\Sigma_1 \leq \Sigma_2$:

$$\begin{aligned}
&\int \mathbf{P}_{\epsilon}\left(\sum_{\tau<t} \max_{(2)} \eta_{v\epsilon}\left(\tau, p, \mu_0, \sigma_1\right) < \pi\right) \mathbb{J}\left(d\tau, dp\right) \\
\leq &\int \mathbf{P}_{\epsilon}\left(\sum_{\tau<t} \max_{(2)} \eta_{v\epsilon}\left(\tau, p, \mu_0, \sigma_2\right) < \pi\right) \mathbb{J}\left(d\tau, dp\right),
\end{aligned}$$

if $\sigma_1 \in \Sigma_1$ and $\sigma_2 \in \Sigma_2$. The integral is taken over a Poisson measure $\mathbb{J}\left(dt, dp\right)$ which means that we integrate the distribution of the second highest bid over the entire price path. If we defined the kernel function $K_{\pi}\left(\cdot\right)$ by the integral

$$\int \mathbf{P}_{\epsilon}\left(\sum_{\tau<t} \max_{(2)} \eta_{v\epsilon}\left(\tau, p, \mu, \sigma\right) < \pi\right) \mathbb{J}\left(d\tau, dp\right) = K_{\pi}\left(v, \mu, \sigma\right),$$

I have provided a set of assumptions assuring that it will be monotone with respect to its arguments for each π. Moreover, given the specific Poisson measure $\mathbb{J}\left(t, p\right)$ we can compute this kernel function. As a result, we reduce the problem of finding the unknown distribution to the problem of solving the integral equation:

$$\mathbf{P}\left\{p_t < \pi\right\} = \int\int\int K_\pi\left(v,\mu,\sigma\right)dF\left(v\right)dG_\mu\left(\mu\right)dG_\sigma\left(\sigma\right)$$

with a monotone kernel. This is a Volterra-type integral equation for the unknown distribution functions, which is known to have a unique solution. This suggests that given the price process and the visibility parameter we can reproduce the distribution of the valuations and the beliefs of the bidders. To find the unknown distributions, I find the subsets of valuations and the parameters of the bidders corresponding to the same cumulative probability of the auction price. To equate the resulting probability distributions of the auction prices, the probability density of beliefs and valuations in the region with a larger kernel should then be smaller than that in the region with a smaller kernel. This allows me to associate ranks with the probability masses in all regions for the distributions under consideration. As the probability masses must sum to one across all the regions, they are automatically determined by the rank of each region.

Finally, consider a shift in the visibility of all auctions by the same constant. Given the monotonicity of the instantaneous demand with respect to the visibility of the auction, this will result in an equivalent shift of the bidder's beliefs. In this case, if the visibility parameter is not fixed, we cannot separately identify the distribution of beliefs and the visibility parameter. Thus the visibility is identified only up to a location shift.

The instantaneous demand function and the size of price jumps which can be estimated from the collection of price trajectories represent the characteristics of the price process averaged over unobserved characteristics of the bidders: valuations, initial beliefs and observation errors. Therefore, they are informative of the conditional moments of the price process - average frequency and average jump magnitude given time and price. Nonparametric identification of such a system of moments can be formally justified using arguments from the identification theory for nonseparable models with an exogenous regressor or an independent instrument (such as in [Che03], [Mat03], and [IN02]). I have shown above that the structure of the simulations from the model is quite similar to nonseparable triangular systems of moments considered in [Che03].

Q.E.D.

1.4 Estimation

To estimate my model I need to solve the optimal dynamic control problem faced by a bidder. This problem can be represented by the differential equation (3.7), which rarely has a closed form solution and it needs to be solved numerically. To simulate the complete equilibrium model, it is necessary to solve a large set of individual bidding problems.

Simulation-based estimation methods used for estimation of structural methods cannot be applied to my continuous-auction models. Examples of these methods are simulated method of moments (including indirect inference as a special case[23]) and two-stage inference (including the hedonic approach[24]). They all require simulations

[23]This method usually referred to as indirect inference (II) was analyzed in [GMR93] and elaborated later to its efficient version adapted to the score estimation method with the auxiliary model generated by the semiparametric estimate of the likelihood of the data, described for example in [GT02] and [GT96].

[24]For instance the two-stage method proved to be useful for estimation of models of imperfect competition as in [BBL07] and [PSD03] of dynamic games as in [HM91] and, auctions as in [GPV00], differentiated product markets as in [BLP95] and [BB02].

from the structural model given the parameters. Such simulations, however, are not feasible in case of my continuous-time model[25]. Therefore, to estimate my structural model I develop a new estimation method that is based on the simulation of the response of the bidders to the observed price process.

The idea of the estimation method is the following. First of all, I can reconstruct the entire continuous price path from the observed actual bid retrospectively. Then for each point of the price path I can simulate the number of entrants into the auction. For each entrant and the incumbent bidders who have entered in the previous time periods, I can compute the optimal response to the price movement. The components of the structural model (entry rate, distributions of valuations and the initial beliefs) can be parametrized or represented non-parametrically. Structural parameters affect the simulated optimal behavior: parametrization of the entry rate affects the simulated number of bidders and parametrizations of distributions of valuations and beliefs affect the individual bids. From simulated entry of bidders and the optimal response for each bidder, we can compute the second highest bid for the specified group of bidders given the structural parameters, which I will call the response of the structural model to the price. In equilibrium, this response should coincide with the price. As the process of entry is stochastic, the distribution of this response should coincide with the distribution of the price. The "true" parameter of the structural model is estimated by minimizing the distance between the distribution of the price and the distribution of the response of the structural model[26].

---

[25]The simulated methods of moments relies on the specific choice moment equations, which is unclear in the model under consideration. The two-stage methods have not been developed for continuous-time models and they can have extremely large standard errors.

[26]The idea of matching distributions for parameter estimation by indirect inference is discussed in [Gal03], where the author suggests using $\mathbf{L}^2$ distance between distributions.

Let me now describe the estimation algorithm in more detail. While most existing empirical research on auctions uses individual bids as observations, a unique feature of this paper is that I consider the entire price path in the auction as the unit of observation. This feature captures the continuous-time structure of the model, where the equilibrium is described by the stochastic behavior of price.

The estimation procedure has three components. In the first component I estimate the joint distribution of the data. The auction proceeds from time 0 to time T, and at each t let p_t be the observed price and N_t be the number of bidders. Let $f(p_t,\, N_t,\, t,\, \gamma_0)$ denote the joint distribution of the object price and the time when the price jumps[27]. This distribution is characterized by a structural parameter $\gamma_0 \in \Gamma$. Since it is possible to observe prices of multiple auctions, it is possible to record several price trajectories and estimate the marginal distribution of price p_t and jump time t using a kernel estimator:

$$\widehat{f}\,(p,\, t) = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{h_p\, h_t}\sum_{i=1}^{I_k}\kappa\left(\frac{p_i^{(k)} - p}{h_p}\right)\kappa\left(\frac{t_i^{(k)} - t}{h_t}\right), \tag{1.11}$$

where n is the number of observed auctions, k is the index of an auction, I_k is the number of price jumps in the auction k, $\kappa(\cdot)$ is a kernel function, and h_p and h_t are bandwidth parameters. Conditions to ensure consistency and asymptotic normality of the estimates are provided in Chapter 3 and impose several restrictions on the kernel function and bandwidth parameters[28].

---

[27] I assume that the joint distribution of realizations of stochastic processes t, p_t and N_t exists. Formal conditions can be found in [GS79].

[28] These restrictions specify that the density that I am trying to estimate actually exists, the kernel function is very smooth and decreases at a fast rate to suppress the influence of "outliers", and bandwidth parameters go to zero as the sample size increases so that there is no asymptotic bias in the estimates.

A problem for estimation in a continuous-time auction is that the movement of the price depends on the entry of bidders, which is considered latent. If there is a large number of simultaneous auctions (so that entry of bidders to the auction can be considered independent across auctions), the kernel estimator will still yield the correct estimate of the marginal distribution of prices and the moments of price jumps. Formally this implies that, under the aforementioned restrictions:

$$\widehat{f}\,(p,\,t) \overset{n\to\infty}{\longrightarrow} \int f(p,\,N,\,t,\,\gamma_0)\,dN,$$

where N denotes the total latent number of bidders who have entered into the auction up to time t. In the technical companion Chapter 3, I gave conditions under which the obtained density estimates are pointwise asymptotically normal:

$$\sqrt{nh_t h_p}\left(\widehat{f}(p,t) - f(p,t)\right) \xrightarrow{\;d\;} N\left(0,\; f(p,t)\left\{\int_0^{\infty} \kappa^2\left(\psi\right)\, d\psi\right\}^2\right). \tag{1.12}$$

This is an important result establishing that even in the continuous-time case, the density estimate possesses the property of asymptotic normality.

The second component of the estimation procedure is concerned with estimation of the distribution of the response of the model given structural parameters and the data. First, simulate the entry of bidders N_t given parameter vector γ for every instant t. Given the structural parameter γ and the optimal bidding problem for each price, I can then calculate the second highest bid in the auction at any given instant. This determines the "response" of the structural model to the data:

$$\widehat{p}_i^{(k)}(\gamma) = \varphi_{\widehat{N},\gamma}\left(p_i^{(k)}\right).$$

Given the price process and the simulated optimal response $\widehat{p}\,(\gamma)$ for every t given parameter vector γ we can estimate the density of the model response as:

$$\widehat{f}_\gamma\,(p,\,t) = \frac{1}{n}\sum_{k=1}^{n}\frac{1}{h_t h_p}\sum_{i=1}^{I_k}\kappa\left(\frac{\widehat{p}_i^{(k)}(\gamma)-p}{h_p}\right)\kappa\left(\frac{t_i^{(k)}-t}{h_t}\right).$$

The structure of this estimator is the same as the structure of the estimator for the density of the data but the observed price is substituted by the simulated price given the parameter vector γ. By construction, the total number of bidders who have entered into the auction $\widehat{N}$ is independent across auctions. This suggests that:

$$\widehat{f}_\gamma\,(p,\,t) \xrightarrow{p} \int f(p_t,\,N,\,t,\,\gamma)dN,$$

where convergence is justified by the same arguments as in (C.3).

The measure distance between the observed joint distribution of the observed and the simulated price and time of the price jumps is chosen to be a Kullback-Leibler information criterion (KLIC), which can be used both for estimation and model selection[29]. The idea now will be to compare the empirical model with the structural model (based on the simulated response of bidders). At the point when the parameter of the structural model is $\gamma = \gamma_0$, both models should work equally well.

Let I_k be the number of price jumps in the auction k and n be the total number

[29] [Vuo89] and [CHS05] show that the KLIC is a powerful tool for model selection.

of auctions. The Kullback - Leibler Information Criterion takes the form:

$$\widehat{\text{KLIC}} = \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{I_k}\log\left[\frac{\widehat{f}(p_i^{(k)},\, t_i^{(k)})}{\widehat{f}_\gamma(p_i^{(k)},\, t_i^{(k)})}\right] \tag{1.13}$$

Minimizing the KLIC is equivalent to maximizing

$$L_n\left(\gamma\right) = \frac{1}{n}\sum_{k=1}^{n}\sum_{i=1}^{I_k}\log\left[\widehat{f}_\gamma(p_i^{(k)},\, t_i^{(k)})\right],$$

which represents the pseudo-likelihood of the model. To compute this pseudo-likelihood we don't need to estimate the density of the observed data. The latter is needed to make specification testing of the model. In Chapter 3, it is shown that the estimate for the structural parameter γ obtained from minimization of the KLIC function will be asymptotically normal, so that:

$$\sqrt{nh_th_p}\left(\widehat{\gamma}-\gamma_0\right) \xrightarrow{d} N\left(0, Q^{-1}\,\Omega\, Q^{-1}\right), \tag{1.14}$$

where

$$Q = E\left\{\int_0^T \frac{\partial f(p_\tau,\tau,\gamma_0)/\partial\gamma}{f(p_\tau,\tau,\gamma_0)}\, dJ(\tau,p_\tau)\right\} \quad \text{and} \quad \Omega = 2\left(\int_0^\infty \kappa^2(\psi)\, d\psi\right)^2 E\left\{\int_0^T \frac{dJ(\tau,p_\tau)}{f(p_\tau,\tau,\gamma_0)}\right\}.$$

A significant problem that might arise in this context is that optimization of (3.8) requires simulating the model response $\widehat{p}$ for each parameter value γ. For especially large models, this simulation can be extremely time consuming. However, the computational burden can be reduced substantially by using Bayesian methods for simulation. In fact, unlike deterministic optimization procedures, a Bayesian estimation procedure utilizes all of the intermediate output (such as function and parameter values) to form the posterior distribution[30].

Let me first describe the procedure for estimation through Markov Chain Monte Carlo (MCMC) and then characterize the asymptotic properties of the obtained estimates. First, we obtain a non-parametric estimate of the density of the data $\left\{ \left(p_i^{(k)}, t_i^{(k)} \right)_{i=1}^{I_k} \right\}_{k=1}^{n}$. Then we construct a random walk sampler from the density $\varphi(\gamma) = \zeta \, \exp\left\{ -\widehat{KLIC}\,(\gamma) \right\}$ for some normalizing constant ζ, where $\widehat{KLIC}$ refers to the data-driven estimate of the KLIC. The random walk sampler is paired with the Metropolis-Hastings method for sampling from an arbitrary density. The method is performed as follows.

The algorithm is initialized at some parameter value $\gamma^{(0)}$. One then proposes a parameter vector γ drawn from the auxiliary distribution. The constraints on the parameter support are easily implementable by a particular choice of a prior distribution. If the KLIC function decreases at the new parameter value, the draw is "accepted" (i.e. we add it to the database of the draws from $\varphi(\gamma)$) with probability $p = \exp\left\{ \widehat{KLIC}(\gamma^{(0)}) - \widehat{KLIC}(\gamma) \right\}$ and the new draw $\gamma^{(1)} = \gamma$. Otherwise, with probability $1 - p$ the draw will be rejected and $\gamma^{(1)} = \gamma^{(0)}$. This process is repeated sufficiently many times to achieve convergence (meaning that the Markov chain of γ becomes stable)[31]. Markov chains generated by this algorithm will be stationary

---

[30] [CH04] have shown that for specific classes of loss functions in minimum distance - type procedures, the first and second moments of posterior distribution formed from the minimum distance criterion have the same asymptotic properties as the maximizer to the minimum distance criterion and its variance.

[31] The fact that the resulting values of γ will represent the distribution with the density $\varphi(\cdot)$ is possible to establish in the following way. Let $\gamma_0 \sim \varphi(\gamma)$. Then acceptance of γ is justified by the fact that $\exp\left\{ \widehat{KLIC}(\gamma_0) - \widehat{KLIC}(\gamma) \right\} < z$. The probability of this event:

$$\mathbf{P}\left\{ \gamma \text{ is accepted} | \gamma_0 \right\} = \int_0^1 \mathbf{1}\left\{ e^{\left\{ \widehat{KLIC}(\gamma_0) - \widehat{KLIC}(\gamma) \right\}} < z \right\} dz = \exp\left\{ \widehat{KLIC}(\gamma_0) - \widehat{KLIC}(\gamma) \right\}$$

under conditions described in [RC06]. These conditions are satisfied for posteriors generated by distance functions with a single minimum and a distance measure increasing at a high enough rate.

[CH04] show that the asymptotic behavior of the posterior mean and variance is the same as the asymptotic behavior of the estimates implied by the minimum distance function. In Chapter 3 it is shown that the asymptotic behavior of the MCMC estimate of γ under a fixed, non-parametric estimation of the data density is:

$$\sqrt{nh_t h_p}\,(\widehat{\gamma}_{MCMC} - \gamma_0) \xrightarrow{d} N\left(0, \frac{1}{2}Q^{-1}\,\Omega\,Q^{-1}\right) \tag{1.15}$$

As compared to the asymptotic variance, the variance of the MCMC estimate has an additional factor of $\frac{1}{2}$. This occurs because the variance of the non-parametric estimate of the data density is not taken into account in the MCMC procedure. To obtain the correct variance estimate, we need to double the MCMC variance estimate.

Although the form of the variance matrix for the sampled parameters is non-standard, it does not imply that the "mean-squared Hessian" is not equal to the variance of the "mean-squared score" of the objective function. This result is presented formally in Chapter 3. Here I will give a simple argument providing intuition for such result. The pseudo-likelihood for the model can be represented as:

$$L_n(\gamma) = \frac{1}{n}\sum_{k=1}^{n}\int_{0}^{T} \log\left[\widehat{f}_\gamma(p^{(k)},\, t^{(k)})\right]\, dJ\left(p^{(k)},\, t^{(k)}\right).$$

The unconditional density is thus expressed as: $f(\gamma) = \mathbf{P}\{\gamma \text{ is accepted}|\gamma_0\}\,\varphi(\gamma_0) = \zeta\, e^{\left\{-\widehat{KLIC}(\gamma)\right\}}$, which is exactly the density $\varphi(\gamma)$. This implies that γ is a draw from this density.

To derive the asymptotic behavior of the estimate $\widehat{\gamma}$ we need to consider the difference between $L_n(\gamma)$ and its population value for the true structural parameter:

$$E\left\{L\left(\gamma_0\right)\right\} = E\left\{\int\limits_0^T \log\left[f_{\gamma_0}(p,\,t)\right]\,dJ\left(p,\,t\right)\right\}.$$

The difference between these values can be represented as a sum of three error components: the component due to the deviation of the estimated parameter from the true one, the component due to the deviation of the estimated density from the true one and the sampling error in the stochastic integral. The last component has a smaller order that the first two ones[32]. The second order components will have similar structure as the first order components. In fact, if γ changes then the estimated density $\widehat{f}_\gamma(\cdot)$ changes as well. Thus, in the second order term the error associated with the density estimation will dominate. Therefore, the expected second-order component will have the same structure as the variance of the first-order component.

1.5 Results of structural estimation

In the previous sections I constructed a continuous-time auction model and proved that the model is non-parametrically identified. I also developed a method for flexible estimation of the model using a Markov Chain Monte Carlo approach. In this section I describe a particular model specification I adopt for my dataset and report the empirical results.

To perform inference, I need to estimate the posterior distribution of the structural parameters using a sequence of simulations. However, the computational burden of

[32]This result is valid because the counting integral in the pseudo-likelihood function belongs to a Donsker class as shown in Chapter 3.

the simulation increases with the dimension of the parameter vector, requiring more evaluations of the distance function. For this reason, I need to specify the structural model parsimoniously enough to allow the chain to explore the posterior parameter distribution in an efficient manner. These parametric assumptions are made for computational convenience and are not required for model identification.

The baseline structural functions describing the dynamics of the price are the instantaneous demand and the size of price jumps. The specification of the instantaneous demand as observed by the individual bidder (given the observation error ϵ) is assumed to have a logistic form:

$$\lambda_\epsilon(t,\, p,\, \theta) = \alpha_\lambda\, \epsilon \exp{(\theta)}\, \frac{exp\left(\mathcal{P}(t,p)\right)}{1+exp\left(\mathcal{P}(t,p)\right)} = \alpha_\lambda\, \epsilon e^\theta \mathcal{L}\left(\mathcal{P}(t,p)\right), \tag{1.16}$$

where $\mathcal{P}(t,p)$ is a polynomial in time and price, $\mathcal{L}(\cdot)$ is the logistic function, and ϵ is uniformly distributed on [0, 1]. The logistic structure of the frequency is not only convenient but also ensures that the frequency stays reasonably bounded[33]. Flexibility in the structure is captured by a polynomial form in the argument of the logistic function.

The magnitude of the price jumps is generated by the logistic form, similar to that for the instantaneous demand function

$$h(t,p,\eta) = \alpha_h \mathcal{L}\left(\mathcal{Q}(t,p,\eta)\right), \tag{1.17}$$

where $\mathcal{Q}(\cdot)$ is a polynomial function of time, price and the individual bid increment. This polynomial structure allows a semi-parametric representation of both the in-

[33]For large frequencies the solution to the individual bidding problem becomes unstable.

stantaneous demand and the size of the price jumps.

The non-parametric flexibility of the structural model increases with the power of the polynomials $\mathcal{P}$ and $\mathcal{Q}$. However, this also increases the number of unknown coefficients in the polynomial representations, leading to a decrease in computation speed. For this reason the degrees of polynomials $\mathcal{P}$ and $\mathcal{Q}$ were restricted to 2. The polynomial $\mathcal{P}(\cdot)$ in the instantaneous demand function is thus quadratic in time t and current price p, taking the form:

$$\mathcal{P}(t,\, p) = a_0 + a_1\, p + a_2\, p\, t + a_3\, p^2. \tag{1.18}$$

The price jump size function h is a logistic function of the quadratic function of time t, price p, and the bid increment η:

$$\mathcal{Q}(t,\, p,\, \eta) = b_0 + b_1\, p + b_2\, t + b_3\, \eta\, p + b_4\, \eta + b_5\, \eta^2. \tag{1.19}$$

The parameters characterizing individual bidders in the continuous-time auction model are their valuations and the parameters of their prior beliefs regarding the visibility of the auction. The valuation of each player was assumed to take the form $v = \underline{v} + \sigma_v\, \xi^2$, where ξ is a standard normal random variable. The parameters $\underline{v}$ and σ_v are estimated. Since this implies that the normal variables ξ are independent across bidders, the bidders have private valuations as in the theoretical model.

To simplify the computations of the individual bidding problem, I assume that there are only three types of bidders (the intermediate type includes all bidders except for the bidders with degenerate beliefs), and each entering bidder is exogenously assigned a type with certain probability. The first type of bidders - the non-strategic bidders - bid their valuations immediately after entry. The probability of the first type is denoted Δ_u. The second type of bidders - the learning bidders - have imprecise information about the visibility of the auction. I estimate the parameters of the distribution of the initial beliefs of these bidders - the mean μ_θ and variance σ_θ. The distributions of initial beliefs across bidders were taken to be normal, truncated to be positive. The third type of bidders - the experienced bidders - are assumed to know the visibility of the auction exactly and behave according to the optimal strategy. The probability of the third type is the complement Δ_i.

I then applied my MCMC estimation procedure to data for 1281 eBay auctions of pop-music CDs. The estimation procedure was organized in the following way. For each auction I simulate up to 50 potential bidders and assign to each of them the observation errors drawn from a uniform distribution on $[0, 1]$ and their private values drawn from the distribution of valuations. Then for the proposed values of the structural parameters $\alpha = (\alpha_\lambda, \alpha_h)$, a_i $i = 1, \ldots, 3$, b_i $i = 1, \ldots, 5$, $\underline{v}$, σ_v, Δ_u and Δ_i, I compute the optimal bid for each bidder and identify the second highest bid. For every auction this simulation process is repeated for the entire price path. On the basis of this bid distribution, I compute the value of the KLIC. The KLIC, non-parametrically computed from the simulated path, is then parsed through the Metropolis-Hasting procedure which generates 50,000 quasi-posterior draws after an initial stabilization period. The estimated parameters which are computed from the mean of the MCMC chain together with the standard deviations are reported below.

$$
\begin{array}{lllllll}
\mathcal{P}(t,p)= & 7.3145 & - & 1.7595p & - \ .7334t & + \ 5.3492tp & + \ 2.1066p^2 \\
 & (1.9049) & & (1.1613) & (2.1024) & (2.1919) & (2.4269) \\
\lambda(t,p,\theta)= & 2.7475 & \epsilon e^{\theta}\mathcal{L}(\mathcal{P}) & & & & \\
 & (1.0376) & & & & & \\
\mathcal{Q}(t,p,\eta)= & 2.5286 & + & 4.2515p & + \ 5.1823t & - \ .9114\eta & + \ .3195\eta^2 \\
 & (3.7750) & & (1.8670) & (2.2704) & (1.2372) & (.0701) \\
h\left(t,\ p,\ \eta\right) = & 1.0776 & \mathcal{L}(\mathcal{Q}) & & & & \\
 & (.4990) & & & & &
\end{array}
$$

$$
\begin{array}{llll}
E\left[\mu_\theta\right] = & 0.4733 & \text{var}\left[\mu_\theta\right] = & 0.7846 \\
 & (0.3598) & & (0.1868) \\
 & & & \\
E\left[\sigma_\theta\right] = & 5.8977 & \text{var}\left[\sigma_\theta\right] = & 1.6799 \\
 & (2.8188) & & (2.9275) \\
 & & & \\
\Delta_i = & .2981 & \Delta_u = & .1116 \\
 & (.0420) & & (.0900) \\
 & & & \\
\underline{v} = & .5385 & \sigma_v = & 4.5286 \\
 & (1.7838) & & (1.6418)
\end{array}
$$

Consider several features of these structural estimates. The first equation, which shows the estimated parameters for the polynomial arguments of the instantaneous demand, suggests that the instantaneous demand is price-dependent. The coefficient for interaction between price and time is positive and significant, implying that elasticity of instantaneous demand is increasing over time. This suggests that the incentive to bid to deter entry changes along the price path.

The second equation describes the instantaneous demand function in terms of the "scale" parameter α_λ. This parameter can be interpreted as the average expected number of bidders who enter the auction. The estimated parameter is equal to 2.74, consistent with the observed bidding patterns since the average number of bidders in an auction is 2.02 in the data.

The third equation represents the polynomial in the price jump size function. The coefficients for time and price are positive, implying as the price increases and time approaches the end of the auction, price jumps approach to their "cap" values. By logistic construction of price jump function, the "cap" value is determined by the factor in the fourth equation. Most of the influence of individual bid increments on price jumps is captured by the last quadratic term. If price and time are fixed, then the influence of the individual bids increases with the size of the individual bid increments until it approaches its "cap". This is reflected in the negative coefficient on the quadratic term of the individual bid increment.

The last four values are the parameters of the distribution of valuations and the shares of the first two types of bidders. The parameter for the distribution of valuations suggests that the lower bound of the support of valuations is statistically indistinguishable from zero. This also implies that the expected valuation of the bidders is equal to \$9.04 (from the mean of the χ^2 distribution). The average price in the auctions is \$6.08. This implies that the bidders are left with a surplus of approximately 30% of the final price. Another key parameter is the proportion of non-strategic bidders Δ_u. The value of the estimated proportion and the corresponding standard error suggest that I cannot reject the hypothesis that the actual number of non-strategic bidders is zero. From parameter Δ_i the proportion of "informed" bidders is 29.81% and is statistically significant. Given these results, I can claim that the estimates of the structural model suggest that the majority of bidders are in fact behaving strategically. The first type bidders are basically naive and they are not behaving strategically. The estimates show that there are very very few of such bidders. The second type bidders are strategic but not experienced enough to have precise information about the visibility of the auction. The estimates show that roughly 2/3 of all bidders are that type. The bidders of the third type are both experienced and strategic and the estimates show that the bidders of this type constitute a large proportion of the bidders population.

1.5.1 Counterfactual simulations of optimal behavior

I now use the parameter estimates obtained in the previous section to demonstrate the properties of the model. I also illustrate how bidding behavior changes in response to exogenous changes in the determinants of the rate entry into the auction and bidder beliefs (i.e., comparative statics). The results of the analysis in this section quantify the effect of the parameters of the model (specifically, the parameters of the entry rate and the bidder beliefs) on the timing and sizes of bids.

Two kinds of aggressive bidding behavior are present in my model. The first is entry deterrence, when bidders raise the price in the auction early to prevent the entry of other bidders. The second is learning prevention, when bidders who are more informed about the unobserved visibility of the auction tend to bid late to hide their information. These two types of aggressive behavior are reflected by the timing of bids. I will analyze how changes in the structure of the model affect the distribution of timing and sizes of bids (relative to the final price).

For the purpose of my counterfactual simulations it is sufficient to look at the behavior of a single bidder because the behavior of other bidders in the auction is fully described by the Poisson price jumps. In my simulation procedure I first produce a "representative" sample of bidders (in terms of their valuations and beliefs) and record the fraction of early bids submitted by a bidder as well as the size of her bids (as proportion of the final price)[34]. This procedure consistently describes the characteristics of the model by looking at the features of their bidding[35]. I simulate the model as described above at the estimated and the alternative parameter values which allows me to find how the bidding behavior is affected by the structure of the model. In the subsequent discussion, I analyze the model by changing the appropriate coefficients proportional to their values at the benchmark (estimated parameter values).

The first feature of interest is the motive for learning prevention. In my model the variance of the bidder beliefs plays an important role in the incentive to prevent learning of other bidders. On the one hand, a bidder who has a large variance of beliefs

[34]For convenience of exposition I define an early bid as a bid submitted during the first 60% of the auction duration. Additional simulations, however, show that the results are not qualitatively different if the timing of early bidding moves up to 95% of the auction duration.

[35]A concrete implementation of the counterfactual simulations will be organized in the following way. (i) choose specific structural parameters which identify the model. The benchmark model is the model with the parameters obtained from the structural estimation. (ii) given the structural parameters draw the characteristics of the bidder: valuation, mean and variance of the initial beliefs about the visibility of the auction, and observation delay error. The distributions of these characteristics are defined by the structural parameters. (iii) for fixed characteristics of the bidder compute the optimal bidding function. The optimal bidding function reflects the value of the optimal bid given time, price and current beliefs about the visibility of the auction. (iv) simulate the price behavior taking into account the optimal bidding function. This allows me to record the timing and the values of bids of the bidder of interest. I repeat steps one to four sufficiently many times (in this analysis - 1000 times) to build a representative sample of the bidders.

engages in the process of updating her beliefs from the price observations referred to as strategic learning. On the other hand, the behavior of bidders with small variance of beliefs is associated with learning prevention. In the first experiment, I simulate the model using different values for the variance of the initial beliefs of the bidders. Table A.3 shows changes in early bidding behavior and sizes of bids if the average variance of bidders' beliefs is set to 0.5, 1.5 and 2.0 times the benchmark value. If the average variance of the bidder beliefs increases, the fraction of early bids tends to increase and bid sizes tend to diminish. This implies that bids become more frequent and more incremental in the presence of strategic learning if the variance of initial beliefs is high. On the other hand, early and frequent bidding becomes rare due to learning prevention if the beliefs of a representative bidder are precise. Table A.3 shows that strategic learning and learning prevention are important even if less than 20% of the bidders belong to the learning type.

Entry deterrence behavior in my model is determined by the price sensitivity of the instantaneous demand, which is characterized by the frequency of the Poisson entry process $\lambda(\cdot)$. The scale factor α_λ and the coefficient a_1 for the linear price term in the quadratic polynomial are used to control for the price sensitivity of the entry process in my model (see equations (1.16) and (1.18)). For both coefficients an increase in either of the coefficients leads to an increase in the marginal entry rate. Table A.3 shows the simulation results, when α and a_1 are set to 0.5, 1.5 and 2.0 times the benchmark values respectively. In general, increases in both coefficients lead to increases in early bidding and the average sizes of bids. The numbers, however, show that the response to changes in the scale factor α is larger. The results of simulations confirm the theoretical prediction that an increase in the sensitivity of entry with respect to price increases leads to more frequent early bidding to prevent entry of other bidders.

Simulation analysis of entry deterrence behavior reported in Table A.3 produces an important testable implication of the model. Consider observations on two markets with different price sensitivities but otherwise similar. In this case we should expect more early bidding in this market. In general, the entry rate reflects the search equilibrium in the auction market, partly determined by the set of auctions that are available to bidders. Taking the number of bidders as given, an increase in the number of choices translates into an increase in the probability of entry into a specific auction. In addition, a high price in the auction negatively affects the surplus of the entering bidders. Therefore the probability of entry into the auction with a high price decreases if the number of choices grows. As a result, the sensitivity of entry with respect to price decreases if the the number of available auctions increases and the average number of bidders in the market remains the same. One way of empirically testing this pattern is to analyze the variation in the early bidding behavior across different categories of items on eBay with different thickness of markets, controlling for the number of bidders. Ceteris paribus the frequency of early bidding should be positively correlated with the number of listed auctions. Another way of testing this implication for early bidding is to run a field experiment where the number of auctions in the market is increased for a period of time. The simulation results imply that we should observe an increase in early bidding behavior in these circumstances. The results of such experiment are discussed in the next section.

1.6 Bidding for musical CDs on eBay: A Field Experiment

1.6.1 Methodology

According to my model, the level of entry deterrence should depend on the sensitivity of the entry rate to price. In fact, the structural estimates presented above indicate that the entry rate is a decreasing function of price. The field experiment described in this section will allow me to exogenously shift the entry rate and change the entry deterrence behavior, providing a strong test of the underlying theoretical model.

Using structural estimation, I found that the rate of entry into the auction decreases as the price grows. I also found that a significant portion (up to 70%) of bidders are the "learning" type - they update their beliefs by observing the evolution of the auction price. These findings suggest that if the bidders are behaving optimally (from the point of view of the model), then we should be able to observe the predicted features of bidding.

Because entry into the auction depends on price, bidders can deter entry by bidding early. Therefore, if it is possible to change the structure of this dependence, we should be able to observe a change in this early bidding behavior. In particular, the rate of entry should depend on the size of the market. If the market is increased (i.e. the number of listed items increases), then the probability of entry into auctions with high prices should decrease, since the chance of winning the item in an auction with no other bidders is higher. As a result, we should expect that early bidding to become more attractive.

The incentive for entry deterrence may also vary across bidders due to differences in their experience. Past bidding experience allows the bidders to predict the entry rate into the auction by forecasting the visibility of the auction. This has different implications for bidders with different prior experience. On the one hand, less experienced bidders will want to "experiment" with the auction to increase their information regarding visibility. On the other hand, more experienced bidders will have an incentive to submit bids only at the end of the auction to conceal their information. For this reason, bidders with different prior experience should respond differently to the exogenous changes in the market, including as an increase in market size. As I argued before, an increase in market size increases the incentive to bid early. If the bidder is experienced, but does not bid early in the market of original size, she may start bidding early if the market expands. As a result, an increase in market size should increase the probability of early bidding by more experienced bidders. Since most of the bidders appear to be experienced, the change in bidding behavior should be substantial.

Field experiments on eBay have been used by several applied researchers to study bidding in Internet auctions. Examples include numerous studies analyzing static models of eBay bidding behavior, such as the effects of seller's reputation ([MA02])), revenue equivalence ([LR99]), and reserve prices ([Rei06]). This is the first paper to study the implications of entry and experience on dynamic bidding behavior. The experiment presented in this section has a unique feature in that it studies the implications of dynamic bidding behavior on early bidding. I perform the experiment by exogenously changing the number of items listed in a particular category on eBay, and observing the change in early bidding behavior.

For my experiment, I chose a specific small item category on eBay and increased the total number of items in this category by listing additional auctions. I formed a control sample of auctions by observing the market prior to treatment. To form a treatment sample of auctions, I listed enough additional auctions to double the size of the market relative to the control group. According to my model, if the treatment increases the sensitivity of entry to price (the elasticity of instantaneous demand increases) then we should observe the entry deterrence feature of bidding behavior discussed above. Specifically, the probability of early bidding across all types of bidders should increase, but it should grow more for more experienced bidders.

In my experiment, I look at bidding on eBay for the Robbie Williams' CD "The Greatest Hits", which was released in 2004. The supply for this CD is quite stable, with an average of 24 items listed each day. This number of listed items allows me to shift the supply significantly by listing additional items. On the other hand, the number of items is sufficiently large to yield a dataset of about two hundred items over a period of two months. I form the dataset by first observing the unperturbed market. Then I double the size of the market and look at the change in bidding behavior.

I next analyze the frequencies of early jump bidding in the treatment (i.e., post-experiment) and control (pre-experiment) groups. According to the theoretical model, the frequency of early jump bidding should increase with the size of the market. I construct a dummy variable for each auction equal to 1 if there is an early jump bid and zero otherwise. Conditional on the auction characteristics (such as seller's feedback score and the location of the seller), we should expect a positive relationship between this dummy variable and the treatment group dummy.

The second relationship studies the dependence between multiple bidding and the bidder's experience. The model predicts that less experienced bidders should bid more frequently than more experienced bidders. Moreover, if the instantaneous demand is low, then experienced bidders will bid only at the last moment of the auction. If the instantaneous demand grows, then the experienced bidders should start submitting early jump bids to prevent entry by potential rivals. The incentive to prevent entry of other bidders offsets the incentive to prevent learning by other bidders. As a result, in the regression across bidders of the number of early bids submitted by a specific bidder on the treatment dummy, experience of the bidder and the interaction of the treatment dummy with experience indicator (conditional on the auction characteristics), there should be a positive coefficient for the interaction term, and a negative coefficient on experience.

1.6.2 Control dataset

The control dataset in my study is a set of auctions for the same CDs before the market was inflated. In total, I collected data for 136 auctions, the earliest auction starting on August 23, 2006 and the latest ending on October 4, 2006. 15 auctions had items located in North America, 69 in Europe (predominantly the UK), 43 in Asia (mainly China and Taiwan), and 9 in Australia. The shipping cost can be substantial if the item is located very far from the winning bidder. In these circumstances, the shipping costs might play an important role in bidding behavior and I will use regional indicators to capture this effect in the data.

I collected the following auction-specific characteristics: buy-it-now price, starting bid, picture dummy, duration (in days), a dummy equal to 1 if the item is available only in the country of origin (regional auction dummy), shipping cost, the percentage of positive feedback for the seller, seller's feedback score, and a dummy equal to 1 if the seller has a store. One minor issue with the prices and costs is that the auctions in the UK and Australia use, British pounds and Australian dollars respectively. All prices and shipping costs were converted to US dollars on the basis of daily average FOREX exchange rate on the day of transaction. Since there were no significant fluctuations of the US dollar exchange rate to these currencies, this conversion should not significantly impact the results. Table A.4 presents summary statistics for the variables in the control sample. Note that the highest buy-it-now prices and the starting bids correspond to rare autographed CDs which might be considered collectibles. These were included in the dataset because they still possess the basic properties of the standard CDs. Most of the items in the sample have a picture. It is worth noting that, in most cases, this is a picture that is offered by eBay automatically and corresponds to the cover of the album [36]. The average duration of the auction is 7 days, with a range of 3 to 10 days[37]. While 70% of CDs in the sample are new, most of the used CDs were auctioned from the UK (40 in total) and none of them were auctioned from Asia. The average shipping cost (consisting of the cost of postage and a handling fee) is close to the average sale price of the CD. This makes the total highest prices of the item on eBay close to the lowest prices available at online retail stores (such as Amazon.com). The largest shipping costs are for the farthest distance between seller and buyer. The shipping cost shown in the table corresponds to the shipping cost to the United States, while the empirical analysis uses the bidder-specific shipping cost. The feedback variable does not seem to be very informative as a majority of sellers in the sample have a 100% positive feedback, probably due to the existing eBay culture of leaving negative feedback only in the case of a very bad experience. A more informative measure is the feedback *score*, which indicates the experience of the eBay participant and is proportional to the number of transactions made by the user on eBay. Most of the sellers in the sam-

[36]This picture and the list of songs can be generated automatically on eBay website if the seller uses the search option allowing to track the album information from the bar code on the cover of the disc. For this reason,it is possibly an unimportant factor for the bidders.

[37]This can be explained by the fact that there is no additional listing fee on eBay for a 7-day listing (and if the item was not sold in can also be relisted without extra charge), while a 10 day listing costs an extra 40 cents, and a 3 day listing requires a certain level of feedback of the seller.

ple have very high feedback scores, and, as the mean of the store dummy indicates, almost 60% of the sellers have stores. This suggests that most sellers in the sample are quite experienced.

Turning to the characteristics of the bidders in the control sample, Table A.5 summarizes the main parameters of the bidders. First of all, we can see that the feedback scores of the bidders are significantly smaller than the feedback score of the sellers. This implies that the sellers in my sample have more experience on average. The percentage of positive feedback, similar to that of the sellers, is equal to 100% for most of the bidders. The last two variables are the dummy "experience" indicators constructed from the observations of past purchase histories of the bidders. The first is equal to one if a specific bidder has won an auction for any music CD on eBay during the last three months. The second variable is equal to 1 if a bidder has won any auction on eBay within the last three months. The CD "experience" variable shows that, in my sample, 45% of the bidders have won a CD on eBay, while 94% of bidders have won any auction on eBay during the last three months. These variables are informative from the point of view of my theoretical model, since they indicate that the bidders are familiar with the price behavior on eBay during the auction and, thus, can use this information to improve their bidding. The last variable reported in the table is a dummy equal to 1 if the auctioned item and the bidder are located in the same region. As one can see, 62% of bidders preferred to bid on items on the same continent.

1.6.3 Treatment dataset

To construct the treatment dataset, I expanded the market by listing additional items. I used 5 different sellers' accounts and listed from 5 to 10 CDs on each one. This allowed me to double the average number of listed CDs per day in the auction.

During the experiment, I sold 60 music CDs. Some of these CDs had to be re-listed because there were no bids in the auctions. The actual dates of the experiment ran from October 4, 2006 to November 10, 2006. To avoid ”transition effects” in the beginning, and to take into account that I did not have enough CDs at the end to double the market, I only used the data from 20 days of the experiment. I used the data for 156 auctions with the first auction starting on October 8, 2006 and the last auction ending on October 28, 2006. In total, 70 items were located in North America (these were predominantly the items listed from my accounts), 56 in Europe, 25 in Asia, and 5 in Australia.

I collected the same variables for the auctions in the treatment dataset as in the control dataset. Table A.4 presents the statistics for the auction characteristics in the treatment sample. The characteristics of the items in the treatment sample are quite similar to those in the control sample. The duration of the auctions in the treatment sample is close to 7 days. Because all the CDs which I listed were new, the percentage of new CDs is higher in the treatment dataset. There are also fewer auctions which restrict the shipping to the country of item location, because my CDs were available to overseas bidders.

Table A.7 reports the characteristics of the bidders in the treatment sample. The average experience characteristics, the dummies for CD purchases and all purchases of bidders on eBay as well as their feedback scores, are very similar in the treatment and control samples. The number of bidders in the treatment sample is smaller because the treatment sample covers a shorter period of time (20 days as compared to 31 days in the control sample). The only visible difference between the treatment and the control sample is that the fraction of bidders who are located in the same region as the item is larger in the treatment sample. This can be explained by the fact that, before my experiment, many bidders from the U.S. were bidding for items in Europe, predominantly in the U.K.

1.6.4 Analysis of early bidding

Using the data from my experiment, I analyze the correspondence between the predictions of my theoretical model and the results of the experiment. The model predicts that bidding early becomes more attractive to bidders due to the increased size of the market. To capture this effect, I constructed an early jump bid dummy variable. This dummy variable is equal to one if, during the first 80% of the duration of the auction[38], there was a bid which increased the price above a certain threshold. To verify the robustness of my analysis, I set the thresholds of price increases determining the jump bid at 10%, 15%, and 20% of the final price in the auction respectively. For this dummy, I then estimated probit models across auctions using the combined dataset of treatment and control groups. While the dependent variables in these probit models are the jump bid dummies, the independent variables include the treatment effect dummy and a set of auction-specific regressors including availability of picture, duration, availability outside the country of item location, seller's feedback and feedback score, and an indicator that the seller has a store. The results of the probit regressions are presented in Table A.8.

The estimates in Table A.8 suggest that the treatment dummy is significant for every jump bid threshold. Therefore, the results of the experiment support the model's prediction that an increase in supply leads to increased early bidding. This effect is quite significant in absolute terms: the probability of early bidding in auctions with doubled supply is 11% higher than in auctions without this treatment.

[38] I also used other thresholds such as 30% and 50% of the duration but the results were quite similar

According to Table A.8, early bidding is affected by the characteristics of the auctions. Specifically, one can see that including a picture of the CD increases early bidding. On the other hand, longer auctions tend to have early jump bids less frequently than short auctions. The increase in the shipping cost decreases overall bidding and, in this way, decreases early bidding as well. Finally, one can see that early bidding is more likely in auctions where the seller has a store. This last result coincides with my interpretation of the visibility of the auction. Specifically, if eBay stores have many cross-listed items, then auctions via eBay stores are more ”visible”. As a result, we should expect that the instantaneous demand in these auctions to be higher and more price-elastic, meaning that there is a greater incentive to bid early. This is confirmed by the data from my field experiment.

The information about individual bidders, specifically the experience dummies and the feedback scores, allows us to examine the effect of bidding experience on early bidding. My model predicts that bidders have an incentive to deter entry and prevent learning. These incentives have the opposite effect on experienced bidders. The entry deterrence incentive leads bidders to bid early to discourage other bidders from entering the auction. The learning prevention incentive leads bidders to bid late to prevent learning. The relative effect of these two incentives should change during my experiment. We know that the increase in supply makes the instantaneous demand more price - elastic. In this case, for more experienced bidders, the incentive to deter entry should dominate the learning prevention incentive. As a result, we should expect an increase in early bidding by experienced bidders. To analyze the effect of increased supply on bidding, I constructed a variable equal to the number of bids that a single bidder submitted in the first 80% of the auction duration. To measure experience, I used the experience dummy variables described in Section 6.2.

In addition, I also used the feedback score of the bidder, which is approximately equal to the total number of transactions the bidder has had on eBay. I ran OLS regressions of the number of early bids on the experience indicators and interactions between the treatment dummy and the experience indicator (controlling for the bidder location). For a given market size, more experienced bidders bid less frequently than less experienced bidders. My model predicts that while all bidders tend to bid more frequently when the market size increases, the largest increase should occur for more experienced bidders.

Table A.9[39] contains the coefficient estimates for the models with different bidding experience indicators. In all three cases, I find a positive and significant coefficient on the interaction term, confirming the predictions of the theoretical model. Moreover, we can see that for the bidding experience indicators, the coefficient is negative, confirming that more experienced bidders indeed submit early bids less frequently due to the dominance of the learning prevention incentive. Analysis of these models with additional explanatory variables confirms that the results are robust with respect to changes in the model specification.

1.7 Conclusion

While it is not uncommon to use standard second-price auction models to study Internet auctions, the co-existence of multiple simultaneous auctions for similar items together with the information heterogeneity inherent in Internet auctions significantly affects bidding behavior. In this paper, I show that in such an environment two types of aggressive behavior by bidders can occur: entry deterrence and learning prevention. To deter rival entry, bidders submit early aggressive bids to raise the price in the

[39]Feedback score variable was normalized by 10,000 to obtain reasonable coefficient values

auction, thereby diverting the entry of rivals to alternative similar items. To prevent their rivals from learning the true "visibility" of the auction, informed bidders will bid late. The observed bidding behavior arises from the interplay between these two forces.

In this paper, I develop a methodology for modeling auctions in continuous-time, allowing for both endogeneity and uncertainty in the entry process, two central characteristics of on-line auctions. I estimate the model using data from auctions for pop-music CDs, validating my theoretical model with empirical data. As an independent test of the model, I conduct a field experiment on eBay. The experiment tests the prediction of my model that, in thicker markets (with more items per bidder), entry deterrence should be more frequent. I increased the number of listed items in the market for a particular music CD and found a significant increase in early bidding. This result both validates my dynamic model of endogenous entry and uncertainty, and also distinguishes it from alternative models that lack similar predictions.

Chapter 2

Empirical Content of a Continuous-Time Principal-Agent Model: The Case of the Retail Apparel Industry

2.1 Introduction

Empirical analysis of incentives provided by executive contracts is a practically important issue. In the empirical literature studying incentive contracts, it is a common practice to consider simple dependencies between components of contracts and characteristics of firm performance. In dynamic environments, this approach will yield unbiased predictions if both the firm owner and the manager observe current state variables, while the actions of the manager are unobserved by the owner. In the case where the manager has privately observed dynamic state variables, the optimal contract structure will be non-linear. In this paper, I argue that this result has the following implications. First, even if the owner of the firm uses a sub-optimal linear contract, the relationship between the contract parameters and the observable characteristics of the firm will be non-linear, while these characteristics become effectively endogenous. Second, endogeneity bias in this case cannot be corrected by conventional instrumental variable methods. This bias can lead to substantial errors in predicting the effects of analyzed characteristics on performance of the manager and, if these estimates are used for contract design, to sub-optimal contract structure.

To convey these arguments, I develop a continuous-time principal agent model with a dynamic state variable which is privately observed by the manager. In my model, motivated by managerial behavior in the retail apparel industry, I consider a single principal (a representative shareholder) and a single agent. State variables of the model include consumer demand and consumer tastes. Consumer demand is observed both by the principal and the agent. However, its dynamics are driven by consumer tastes, which are assumed to be observed only to the agent. This induces correlation between demand and consumer tastes even though random shocks in tastes and demand are uncorrelated. Consumer demand is also assumed to depend on the brand characteristic of the firm. The manager optimally sets this characteristic, exerting effort and decreasing her utility. The principal uses a linear contract which aims at setting incentives for the manager to set the brand characteristic optimally from the point of view of the principal. The optimal strategy of the manager will be a function of tastes and demand. Therefore, given that characteristics of firm's performance depend on consumers' tastes, they cannot be used as exogenous explanatory variables.

In this paper, I develop an estimation methodology which can alleviate this endogeneity problem. The method is based on explicit consideration of the manager's optimization problem and using computed manager's best response to consistently estimate the incentive effects. I provide an empirical strategy to identify the model from the data and estimate its parameters using a non-parametric procedure, and adapt this methodology to a particular industry. The structural estimates can be used for the welfare analysis of both long and short-term effects of changes in legislation in the industries and for long-term predictions of the effect of these changes on the productivity of managers and firms' output. To illustrate my methodology, I use data from the ExecuComp dataset on executive contracts in the retail apparel industry. My structural estimates allow me to evaluate the welfare effect of the introduction of Sarbanes and Oxley act on the apparel retail industry.

The analytical procedure suggested in this paper proceeds in three steps. In the first step, I construct a continuous-time model to describe the manager's response to the contract incentives. I describe the structure of the state variables, information available to the principal and the agent, and the way the agent can influence the dynamics of the state variables. In practice, the analysis of contractual incentives should be industry-specific because the structure of observed and unobserved state variables can vary greatly by the industry, which can result in significant structural differences between contracts in different industries.

In the second step, I use the constructed structural model to design an estimation procedure to recover structural parameters for a particular industry. This step depends on the structure of the ExecuComp dataset which I use for empirical illustration. The estimation method used in the second step combines simulation from the model conditional on actually observed characteristics of the manager and the firm. It is based on comparing the simulated distribution of state variables with the actual distribution. I show the model which I study is non-parametrically identified. Simulating the model in the second step requires computation of the manager's best response. To fulfill this task, I develop a fast and reliable algorithm for solving for the best response of the manager and verify that the noise associated with simulations is infinitesimal as compared to the noise in the data.

In the third step, I illustrate the model by estimating its structural parameters and use the obtained structural estimates to analyze the effect of exogenous industry changes on profits and revenues of the firm under the presence of an agency problem.

For an empirical illustration, I use data for a particular industry: apparel retail. The sales in this industry are driven by volatile consumer's tastes and display a significant amount of seasonality. The volatile structure of market tastes is frequently observable only by the manager (but not by the shareholders). This makes tastes an unobserved state variable. In the subsequent analysis I develop a model which is more general than [HM87] to describe the executive compensation in the apparel retail industry. I estimate the model for ExecuComp data, and use the estimates to predict the effect on the industry due to the introduction of the Sarbanes and Oxley act (which imposed strict institutional restrictions on the actions of executive managers). Although I observe a behavioral response to the incentives provided in managerial contracts, a significant proportion of the compensation in the industry is in the form of fixed base salary. My analysis demonstrates a significant response to the introduction of the Sarbanes and Oxley act which increased legal manager's liability for the financial results of the firm. In particular, I find that the introduction of the act resulted in an increase both in the base salary and bonus parts of managers' compensation and has led to shifting the welfare loss towards the firm owners.

My paper fits broadly into the empirical literature on the economics of asymmetric information and, specifically, structural estimation of models of adverse selection and moral hazard. These issues have been studied, for instance in [ACHPng], which deals with disentangling the effects of moral hazard and adverse selection in an equilibrium market setting. [Laf92] and [LS99] analyze the franchise decisions of the firms dealing with the effects of adverse selection in the market. [MM00] focuses on the cross-industry analysis of moral hazard in managerial compensation. The studies in [LSZ00] and [CSS04] use linear models to estimate effect of contracts' incentives in the financial industry. My study differs substantially from the existing literature. First, from a theoretical perspective, I emphasize the importance of using a flexible dynamic model based on computed best responses to produce unbiased structural estimates. Second, I use a computational algorithm for the manager's best response and do no rely on its reduced-form specification. Third, my procedure can be used to produce counterfactual simulations directly. These features allow me to study dynamic incentives within the context of structural parameter estimation in which I nest the numerical solution for the best response of the manager into the estimation procedure.

My paper analyzes the role of dynamic incentives in contract design when the compensation of the manager's actions depends on her continuous - time performance by filling two gaps in the literature. First, although performance-based compensation for top management has been common for decades, it is unclear how close the structure of payments has been to the optimal contract structures. Second, as the structure of informational rent in the dynamic case is more complicated than that of the classical principal - agent problem, the structure of distribution of welfare associated with the distribution of informational rents has not been investigated in sufficient detail. In this paper, I address both issues and show for an example of contracts in the apparel retail industry, how one can produce the estimates of welfare effects from dynamic risk sharing and analyze how this effects change as the structure of incentive compensations changes.

The structure of this paper is as follows. In the first two sections, I develop a methodology for construction and estimation of a continuous-time model. Section 2 discusses the model and Section 3 discusses identification of the model and estimation technique. Section 4 then applies the developed methodology to estimation of the model for executives in the apparel retail industry. I also demonstrate how the obtained structural estimates can be used to produce the evaluations for welfare effects and changes in the production dynamics as a result of the introduction of the Sarbanes and Oxley act. Section 5 concludes.

2.2 A Continuous-Time Principal-Agent Model

In this section I develop a continuous-time principal-agent model. Using a particular structure for the state variables I specialize the model to the apparel retail industry.

2.2.1 Setup

In this section I present a continuous-time model of a managers's response to the contract incentives in the apparel retail industry. Similarly to the setup in [HM87], I assume that the contract has fixed duration T and has to be renewed each period. The terms of the contract are designed by the owner of the firm (a representative shareholder). I simplify derivations in the analysis by assuming that the manager does not own a significant portion of the firm and thus relies only on compensation rather than on dividends.

The structure of the state variables in my model is significantly different from that in the classical setup of [HM87]. The state variable has two components. One component, which is denoted θ_t, is idiosyncratic and represents consumers' tastes. The other component is the sales of the firm D_t, which the manager can control optimally. The dynamics of market sales is determined by the following equation:

$$dD_t = \psi(t, D_t, x_t - \theta_t)\, dt + \zeta(t, D_t, x_t - \theta_t)\, dB_t^D. \tag{2.1}$$

In this equation, D_t is the demand for the production of the firm at time t, where I assume that the firm does not experience stock-outs. The parameter characterizing the tastes of consumers in the market is denoted by θ_t. The brand characteristic is denoted x_t which reflects the correspondence of the production variety offered by the firm to the variety of products demanded by consumers, and B_t^D is a Brownian motion. This form of demand dynamics reflects the cyclical structure of the ready-to-wear clothing market. The first component describes the deterministic trend driven by consumer tastes. The second component describes random deviations from the deterministic trend. Demand volatility increases if production variety offered by the firm does not reflect market demands.

There are two features which differentiate my model from the existing literature. First, I consider a state variable which is not observed by the principal and represents consumer's tastes. Second, I allow the agent to control both the drift and the diffusion of the state variable. In the presence of both features, the contract structure developed in [HM87] is no longer optimal even in the class of linear contracts. While the first-best optimal contract in the presence of unobserved state variables is typically non-linear, I study optimality within a class of linear contracts which are computationally feasible and do not involve a numerically intractable problem with non-linear contracts. Some arguments in favor of the linear contract structure are given in [OY03], where the agent is also allowed to control the variance of the state variable.

The drift component of the demand $\psi(t, D_t, x_t - \theta_t)$ is a function of time, as well as current demand volume D_t and a brand parameter $x_t \in [0, 1]$, which will be referred to as a "brand characteristic". It is assumed that $\psi(t, D, z)$ is maximized at $z = 0$ for any $t \in [0, T]$ and $D \in \mathbb{R}_+$, that is, the deterministic component of demand is the highest if production fashionability completely reflects the consumer taste parameter. The function $\psi(\cdot)$ reflects seasonal changes in demand.

The diffusion component $\zeta(t, D_t, x_t - \theta_t)$ is a function of the same set of variables that enter the drift component. The dependence of the diffusion term on the demand level reflects possibly higher fluctuations of demand in the firms with high volume of production. The dependence of the diffusion term on the difference between the market tastes for variety and the variety of production offered by the firm captures a probable increase in demand volatility if the brand characteristic of the firm is far from the market tastes.

Market demand in this paper is analyzed as a stochastic process. Such analysis implies a set of assumptions that I am imposing on both the demand dynamics and the industry structure. In particular, I assume that firms in the industry operate on a differentiated product market where each firm faces the demand for its particular market segment. Moreover, I assume that the demand on other market segments does not provide information relevant for predictions of consumers' behavior in a chosen segment. This assumption likely holds for the retail apparel industry, which exhibits a high degree of horizontal differentiation. Each firm in the apparel retail industry offers a mix of consumer products targeted at a particular niche of the consumers' market. I assume that the output of a firm can be expressed in terms of volume of a composite commodity. Moreover, the prices on the market correctly reflect the utility weights of particular products of the composite commodity. Therefore, the dollar value of firm's sales can be interpreted as the product of a price index and a quantity index. The price index reflects the unit value of the composite commodity. The quantity index reflects the quantity of the composite commodity where individual purchases are weighted by the utility indices. The brand characteristic of the firm then will represent the position of the firm in the differentiated product space. Therefore, the shares of purchases by a representative consumer from different firms will be driven by the relative locations of the preferences of the representative consumer and brand characteristics of firms. The demand of consumers for a particular brand is summarized by equation (2.1).

I consider the total sales of the firm rather than the physical volume of sales. While volume data on the apparel retail are available, they are problematic because, apart from the brand policy, the firms also have the price policy which can reflect the quality of the product. In fact the same model of clothes can be made of fabrics of different quality. The production cost and the price can differ significantly even across very similar pieces of clothing. Another important aspect is in-season pricing. In particular, if after the introduction of a new product the firm discovers an unexpectedly high demand for the item, it can temporarily raise the price for this item to avoid stocking out before additional quantity is produced. The data for sales represents the value-weighted demand for the production of the firms and solves this problem. The tastes of consumers then determine the consumer's choice between both the quality and the price dimension of the production of the firm. Specified structure of sales of the firm reflects an additional assumption that consumers are generally price-takers and choose among a menu of objects on the market with different prices and qualities.

The consumers' taste parameter θ_t is assumed to follow a driftless diffusion process:

$$d\theta_t = \gamma(t, \theta_t)\, dB_t^{\theta}. \tag{2.2}$$

The drift term represents a long-term dynamics in the tastes evolution which should not play a significant role in the decisions over a year although it might be important in the long-run. Moreover, the diffusion structure of equation (2.2) significantly simplifies the empirical analysis. The firms are assumed not to influence the dynamics of tastes. This equation suggests that the firm cannot influence consumers' tastes. This assumption is in line with modern research in apparel marketing and consumer psychology (such as, for instance [Ric96], [Spr81], [Spr85]), which suggest that modern apparel production is mostly consumer-driven.

The dynamics of tastes over time are driven by the change of fashion and fluctuations in the quality of output of a specific producer over time. For an individual consumer the movement of tastes changes her marginal utility of a specific brand and affects the demand. Managers of firms are different in their abilities of predicting the future values of θ_t. The market share of the firm at a specific instant will be determined by the closeness of the firm's characteristic x_t to the consumer tastes characteristic θ_t, given the price of the production of the firm relative to the prices of other firms.

There are two sources of asymmetric information in the agency problem of my model. First, I consider publicly traded firms with independent managers, in which there is a separation of ownership and control over the firm. Second, managers act in a volatile market environment and the exact financial outcomes of the firm at end of the contract period are unpredictable. To a large extent, the volatility of the market is caused by the heterogeneity of consumers across the market, and heterogeneity of aggregate consumer's tastes across time. With the pool of potentially heterogeneous consumers, the market share of the firm at a specific time is determined by the price of the firm's output, the profile of tastes of the consumers in the market, and the profile of the stocks of the commodity accumulated by consumers. This consumer taste heterogeneity is associated with the randomness the financial outcome of the firm. A poor financial outcome in a specific period may be either the result of poor brand structuring policy by the manager or low seasonal demand for the product of the firm. The owner of the firm (who characterizes the representative shareholder) wants the manager to choose the best branding strategy to match the consumer tastes and to improve the market sales of the firm. Matching the market tastes is, however, costly to the manager because it requires a significant amount of market research each season to determine the optimal structure of the production mix. The lump-sum initial compensation for the manager will not create sufficient incentives for her to match the market tastes. The owner of the firm needs to design a contract for the manager to increase her effort level. However, the firm owner can only observe the continuous-time sales D_t and use it for contract design. The brand characteristic x_t as well as consumer's tastes θ_t are not observed. The interpretation for this assumption is that: although the shareholders can observe the market sales of the firms from the market reports, to analyze the consumer tastes at a specific part of differentiated product market one needs to undertake costly market research.

I assume that the firm's instantaneous profit is a linear function of total sales $\delta\, D_T$ where D_T is the cumulative sales at the terminal time T. The contract offered to the manager by the owner takes a linear form:

$$S(D) = \sigma(D_T) + \int\limits_0^T \alpha\left(t, D_t\right)\, dt + \int\limits_0^T \beta\left(t, D_t\right)\, dD_t, \tag{2.3}$$

so that the contract is defined by unknown functions $\sigma(\cdot)$, $\alpha(\cdot)$ and $\beta(\cdot)$, while $S(D)$ is paid to the manager at the moment T. Most of the actual contracts used in practice have a linear structure because, first, it is tractable and, second, it is feasible for implementation. Moreover, a non-linear optimal contract schedule might not exist in general[1].

Analyzing linear contracts both simplifies the computations and can approximate

[1]In some cases that the optimal incentive contract might suggest imposing infinite penalties on undesirable outcomes.

well the contract structure used in practice. The lump-sum part of the contract $\sigma(D_T)$ characterizes the salary of the manager. The last two components characterize the incentive payouts (bonuses). The first part of bonus $\int\limits_0^T \alpha\left(t, D_t\right) dt$ sets incentives for the manager to achieve a high level of sales by the end of the contract period. The second part of bonus $\int\limits_0^T \beta\left(t, D_t\right) dD_t$ rewards the growth rate of sales during the contract period. This reward is related to risk aversion and can be justified by the loss aversion of the shareholders, who are concerned that higher growth numbers are better perceived by the stock market. Alternatively, a habit formation argument suggests that the shareholders have preferences for faster growing sales due to positive correlation between current share purchases and future share purchases.

In structure of the contract implied by equation (2.3) that the shareholders do not use the past values of the demand to evaluate the behavior of the manager retroactively. The manager therefore is not "rewarded" or "punished" for the behavior in the previous periods. Because of the diffusion assumption on the consumer tastes dynamics in my model past observations do not provide any basis recovering the future values of tastes. A more complex structure of the consumer tastes could potentially create problems with model identification because the tastes are not observed and I do not have a valid proxy for this variable. For these reasons I do not consider more complicated consumer tastes dynamics.

The manager is assumed to be risk-averse and derives utility from the monetary wealth of the manager minus the cost of effort on the job expressed in monetary terms. The effort of the manager will be called branding in this paper. Branding involves a decision about the production mix during the season and changes in the firm production between seasons. The instantaneous cost of branding is defined by $c\left(x_t - \theta_t\right)$ so that the total cost for the manager is $C(x) = \int_0^T c\left(x_t - \theta_t\right) dt$. The function $c(\cdot)$ is assumed to have the following properties:

- $c(\cdot)$ is symmetric so that $c(-x) = c(x)$,
- $c(x) > 0$ for all $x \in \mathbb{R}$,
- $\frac{\partial c(x)}{\partial x} < 0$ for all $x > 0$.

The branding cost $c(x)$ derives from manager's effort to match the consumer tastes in the market which includes conducting the market research, organizing the work of a design team, making choices about the types and quality of fabrics, and pricing of products during the season to best target the demands of a specific category of consumers. These actions are costly to the manager both in terms of the personal effort and in terms of the expenditures of the firm. In addition, effort can be costly because high production costs not offset by high sales can deteriorate the reputation of the manager. Given the contract structure and cost, the wealth of manager is defined as:

$$W_m = S(D) - C(x). \tag{2.4}$$

The manager is assumed to be risk averse with a utility function

$$U_m(W_m) = -\frac{1}{R_m} \exp\left\{-R_m W_m\right\}. \tag{2.5}$$

The wealth of the owner is the profit of the firm less the value of the contract offered to the manager:

$$W_o = \delta D_T - S(D). \tag{2.6}$$

The owner, similarly to the manager, is assumed to be risk - averse with the utility defined by

$$U_o(W_o) = -\frac{1}{R_o} \exp\left\{-R_o W_o\right\}. \tag{2.7}$$

The owner designs the contract to maximize the owner's utility over the contract space subject to the optimization problem of the manager given that the utility of the manager is not less than the reservation utility.

2.2.2 Manager's problem

The manager is assumed to be interested only in the final payout. Her utility is given by (2.5). The manager solves the problem of utility maximization by finding the optimal branding strategy for the firm given the contract and the dynamics of the demand and the consumer tastes.

The problem of the manager is to maximize her utility during the given period of time such that her wealth is a function of the sales of the firms and cost reflects effort of branding policy. The value function of the manager at time t is defined as the expected cumulative utility up to the end of the period given the information at time t. The value function of the manager at time t can be written as:

$$V(t, \theta_t, D_t) = -\frac{1}{R_m} \sup_{x_t} E_t \left\{ e^{-R_m\left(\sigma(D_T) + \int\limits_t^T [\alpha(t,D_t) - c(x_t - \theta_t)]\,dt + \int\limits_t^T \beta(t,D_t)\,dD_t\right)} \right\}. \tag{2.8}$$

The idea behind the optimization problem of the rational manager is that she will try to build a brand design strategy which will alleviate the non-systematic risk in the dynamics of her utility over time. The problem of the manager has the structure studied in the stochastic dynamic optimal control theory as in [GS79]. I derive the

Bellman equation, which can be written in the form of a partial differential equation:

$$\begin{aligned} &V_t + \sup_{x(\cdot)} \left\{ V_D \left(\psi - R_m \beta \zeta^2\right) + \tfrac{1}{2} V_{DD} \zeta^2 + \tfrac{1}{2} V_{\theta\theta} \gamma^2 - \right. \\ &\left. - V R_m \left[\alpha + \beta\psi - c(x-\theta) - \tfrac{R_m}{2} \beta^2 \zeta^2\right] \right\} = 0. \end{aligned} \tag{2.9}$$

In this equation, subscripts denote the corresponding derivatives of the value function. Equation (2.9) expresses the law of motion for the value function of the manager in terms of its derivatives over the state variable for each instant. The last term in the supremum shows that if the value function of the manager is constant in the state space, then it would be growing exponentially over time. The growth rate will increase if the "instantaneous" compensation of the manager increases and diminish if the costs or the variance of demand shock increases. This growth rate depends the risk aversion of the manager: if her risk aversion is sufficiently small, then an increase in the risk aversion increases the growth rate of value function, if it is very large, then an additional increase leads to the drop in the rate of growth of value function. The first term in the supremum reflects the influence of the marginal value of the manager with respect to the demand. If the demand is increasing over time rapidly enough, then a higher sensitivity of value function implies a lower increase in the value function over time. The second and the third term reflect the effect of the curvature of the value function and variances of shocks in the state variables on the dynamics. In general, a larger variance of shocks in the state variables leads to the slower growth in the value function. This means that the manager in this model "dislikes" uncertainty in the dynamics of state variable.

Intuitively, equation (2.9) can be obtained by equating to zero the systematic drift component in the stochastic Bellman equation for the manager. Informally, one can derive this condition by writing the Bellman equation for expected utility of the manager at time t in terms of utility at time $t + dt$ and optimal actions. Taking the first-order condition, one can derive expression (2.9) using the rules of Itô calculus.

The structure of the optimal control problem suggests that at each instant of time the manager finds an optimal level of effort given the value of the taste parameter and realization of demand. Then she recalculates the optimal policy at the next instant and use it to derive the continuation payoff given by the time derivative of value function. In this way the manager solves for the optimal effort at each value of the taste parameter, demand and time.

In the above derivation of the value function, the manager only maximizes the utility from the monetary payoffs. An implicit assumption here is that the managers do not have other objectives such as the size of the staff of the firm or the long-term value of the firm. This assumption is likely to hold in the industries with a significant control of firm owners over the top management.

2.2.3 Owner's problem

In the linear manager's contract, total compensation is formed by a fixed component and a stochastic component which depends on the dynamics of the firm's sales. In particular the contract structure might not only reward a high level of sales, but it also targets a specific growth rate of sales of the firm. The owner maximizes the expected utility derived from the share of the firm's profit minus the managerial compensation. Similarly to the manager, the owner has a CARA utility defined in (2.7). The owner's problem is to find the optimal parameters of the contract, including the salary and the bonus of the manager's contract, which maximize her expected utility at the end of the contract period[2]. Given the optimal response of the manager

[2]This is based on the assumption that I consider a competitive labor market for executive managers and thus, contracts can be signed again each period. In equilibrium with repeated contract signing

I rewrite the components of the contract to simplify the statement of the optimal contract design problem, that needs to be solved by the shareholders. In [OY03] it has been shown that the optimal linear contract structure is defined by the following expression:

$$\begin{aligned} S(D) = \overline{W}_m + \int_0^T c(x_t - \theta_t)\,dt + \tfrac{R_m}{2}\int_0^T \left(\beta - \tfrac{V_D}{R_m V}\right)^2 \zeta^2\,dt \\ + \int_0^T \left(\beta - \tfrac{V_D}{R_m V}\right)\zeta\,dB_t^D, \end{aligned} \tag{2.10}$$

where $\overline{W}_m$ is the manager's opportunity income. This contract is written in terms of the "effective" reward for demand growth $\beta - \frac{V_D}{R_m V}$. This effective reward will be lower if the value function of the manager is more sensitive to the changes in the market demand. Thus, the compensation will be more skewed towards the fixed salary payout. Such contract structure has a meaningful interpretation. A non-idiosyncratic part of the contract is determined by the cost of the branding effort of the manager and the risk premium for the possibility of having low demand in spite of the high branding effort. This premium is defined by the conditional heteroscedasticity term in the consumer demand, contract compensation for the steady demand growth and the risk aversion of the manager. This formula suggests that more risk averse managers should have a higher compensation for the markets with high demand variance. The idiosyncratic compensation of the manager is determined by the stochastic demand shock and the compensation of the manager for the steady growth of the consumer demand. The optimal contracts depend only on the current but not past values of the consumer tastes parameter. This is due to the lack of path-dependence in

under certain conditions the value function of the owner will will depend only on the parameters of the current period.

the consumer tastes process which is driven by a diffusion process without a drift[3]. Therefore, the contract of the manager is determined only by the demand shocks and not by the taste shocks.

The value function of the firm owner can be written similarly to the value function of the manager in the form:

$$V^*\left(t,\, D_t\right) = -\frac{1}{R_o}\sup_{x,\beta} E_t\left\{e^{-R_o\left(\delta D_T - \overline{W}_m - \int\limits_t^T c(x_t-\theta_t)\,dt - \frac{R_m}{2}\int\limits_t^T \tilde{\beta}^2\zeta^2\,dt - \int\limits_t^T \tilde{\beta}\zeta\,dB_t^D\right)}\right\},$$

where $\tilde{\beta} = \beta - \frac{V_D}{R_m V}$. The equation for the owner's continuation payoff can be derived by applying Itô calculus to $V^*\left(t,\, D_t\right)$:

$$\begin{aligned} V_t^* + \sup_{x(\cdot),\,\tilde{\beta}(\cdot)} &\left\{V_D^*\left(\psi + R_o\zeta^2\right) + \tfrac{1}{2}V_{DD}^*\zeta^2 \right. \\ &\left. + V^*\left[R_o c + \tfrac{R_o}{2}\left(R_o + R_m\right)\tilde{\beta}^2\zeta^2\right]\right\} = 0. \end{aligned} \tag{2.11}$$

The problem of the owner resembles the problem of the manager. The optimal strategy of the owner is similarly computed in two stages. In the first stage, the owner calculates her preferred value of effort of the manager and a contract structure that can make the manager exert effort equal to the desired level. In the second stage, given the time derivative of her value function, the owner recomputes the value function at the next moment of time. This continuous-time backward induction allows me to compute the optimal contact structure along and the target level of effort.

[3]In the presence of fashion cycles one would expect that "tastes" will have a drift. However, if firms in the industry are aware of the existence of these cycles and take them into considerations, I can argue that I only model the responses to deviations from the long-term trend.

One of the main assumptions here is that the firm owner knows the functional form of the demand process and the manager's cost so there is no asymmetric information about the manager's type. In the case where the manager's type is unknown and her cost of effort can vary across types, then the contract structure can be more complicated. Then the contact will take into account additional "truthful reporting" incentive constraints forcing the manager to behave optimally given her type. Comparing the optimal problem of the manager and the firm's owner suggests that a higher cost of branding effort will decrease the growth rate of the value of the manager over time. It will also decrease the growth rate of the value of the firm owner over time, because the effort level of the manager will decrease. A similar effect is induced by risk aversion: a higher risk aversion reduces the growth rate of the value function over time for both the manager and the owner.

Together equations (2.9) and (2.11) describe the structure of the optimal dynamic contract between the principal and the agent. Equation (2.11) defines the optimal contract structure and the target level of effort from the point of view of the firm owner. Equation (2.9) defines manager's effort given the contract structure. As was discussed above, the compensation and the bonus parts of the contract can be described in terms of the manager's cost of the branding effort, the risk aversion parameter and functions describing the drift and the variance of the dynamics of consumers' demand.

2.3 Identification and Estimation

In this section I develop an estimation strategy to estimate a continuous-time principal-agent model developed in the previous section. Before that, I investigate whether the model is non-parametrically identifiable given a sample of these variables of a sufficient size. The identification strategy is specialized for the data that I am using for structural estimation.

2.3.1 Identification

In this section, I demonstrate that the theoretical model of dynamic optimal contract between the manager and the owner is non-parametrically identifiable given the structure of the data. Identification of the theoretical model of the manager's behavior is a complex problem because the strategy of the manager is driven by the latent preference parameter θ_t, which is unobserved both by the shareholders and by the econometrician. In the following I describe the general structure of the data, outline the model assumptions, and discuss their role in the identification of the model given the data structure.

I assume that the information set of the econometrician is similar to that of the firm owner, who observes the demand process but not the consumer's tastes. The econometrician observes the demand realizations at the end of the period D_T and corresponding values of salary $\sigma(D_T)$ and bonus $S(D) - \sigma(D_T)$. The entire path of demand and tastes during the contract period are not observed. The available data consist of the compensations (salaries W_{ik} and bonuses P_{ik}) for a panel of managers $i = 1, \ldots, N$ observed over a collection of time periods $k = 1, \ldots, K$. The available data include the sales of the firms for each period and a set of covariates consisting of the indicators of the firm performance for a specific period of time, and a set of individual-specific characteristics of the managers.

In the model the tastes parameter θ_{ti} plays the role of the unobserved error term in the manager's decision problem and satisfies the following assumption:

Assumption 2.1. *The taste parameter θ_{ti} does not have a drift and is independent across managers $i = 1, \ldots, N$. The compensation of the manager is only a function* *of the firm's sales and does not depend directly on the unobservable tastes.*

This assumption states that each firm is faced with the firm-specific tastes for a product on a differentiated market.

The unobserved dynamics of the tastes translates into the demand on the market for the production of a specific firm and subsequently affects the behavior of the manager. Given the dynamics of the tastes described by the measure of variance $\gamma(\theta,\, t)$[4], the manager solves the utility maximization problem given her risk aversion R_m, cost of effort $c(x_t - \theta_t)$, the parameters of demand $\psi(\theta_t - x_t, t, D_t)$ and $\zeta(\theta_t - x_t, t, D_t)$, and the parameters of the contract $\alpha(t, D_t)$ and $\beta(t, D_t)$. The optimal branding strategy of the manager x can be defined as a functional $x(t, D_t, \theta, c(\cdot), \zeta(\cdot), \psi(\cdot), \alpha(\cdot), \beta(\cdot))$. Note that the continuous-time structure of the model and Markov structure of state variables allow us to avoid considering their dependence from lagged variables. The dynamics of the sales will be a functional of the manager's actions. Therefore, the sales of the firm can be written as $D\,(t, \theta, c(\cdot), \zeta(\cdot), \psi(\cdot), \alpha(\cdot), \beta(\cdot))$. The following identification assumption is important and concerns the parameters of the contract.

Assumption 2.2. *For the set of covariates Z, functions $\alpha(\cdot \mid Z)$ and $\beta(\cdot \mid Z)$ given Z, which determine the wage and the bonus of the manager, are the same for all $i = 1, \ldots, N$*

This suggests that the differences in the specific contracts can be attributed to the observable differences in the structure of the firm and features of managers. Conditional on this observed heterogeneity, this suggests that the contract structure is the same across the firms in the industry.

[4]The variance of tastes at instant t will be determined by the integral of $\gamma(\cdot)$ over time

There are two approaches providing identification of the contract parameters. In the first approach I assume that the contract characteristics are homogeneous of degree 1 in time, and I assume that a large panel of data is available. Then, variation in the compensation across periods allows me to prove identification of contract parameters. In the second approach, I assume that the contracts are optimal from the point of view of the principal. As a result, the parameters of the contract will be functionally dependent on the basic features of the model: demand and cost function. The problem of identification of the contract parameters will reduce to identification of the demand and cost functions.

Consider identification of contract parameters from panel data. Assume that both $\alpha(\cdot)$ and $\beta(\cdot)$ are homogeneous of degree 1 in time. Using observations on the sales and covariates, functions $\alpha(\cdot)$ and $\beta(\cdot)$ can be estimated non-parametrically given covariates. Intuitively, for each data point i and k, I treat the values $\alpha(t_k, D_{ki})$ and $\beta(t_k, D_{ki})$ as unknowns. Moreover, I find the grid points for time as $t_k = kT$. Denote $\tau_k = \frac{k}{K}$. Then, writing the increase in the compensations as:

$$\Delta P_{ik} \approx \alpha(t_k, D_{ki})\Delta t + \beta(t_k, D_{ki})\Delta D_{ki} = \frac{T}{K}\left(\alpha(\tau_k, D_{ki})\Delta t + \beta(\tau_k, D_{ki})\Delta D_{ki}\right)$$

I interpret the result as a system of NK equations with NK unknowns. In these equations $\tau_k \in [0, 1]$ and represents the grid for time in the unit interval. Thus, one can recover the structure of the contract parameters from the data. Note that this result holds for any degree of homogeneity of contract parameters with respect to time.

It is not possible to simultaneously identify the scale of tastes separately from the functions $\psi(\cdot)$ and $\zeta(\cdot)$. Consider the scale transformation $\gamma(t, \theta) \mapsto Q\gamma(t, \theta)$

for $Q > 0$. Then I define the functions $\overline{\psi}(t, \theta - x, D) = \psi(t, (\theta - x)/Q, D)$ and $\overline{\zeta}(t, \theta - x, D) = \zeta(t, (\theta - x)/Q, D)$. The demand equation corresponding to $\overline{\psi}(\cdot)$ and $\overline{\psi}(\cdot)$ remains exactly the same as before the transformation. A similar argument applies to a constant shift in the taste parameter. An additional scale assumption should be imposed on the time argument. Note that as the demand is observed only as a cumulative quantity at the end of the auction, both scale and shift transformation change the shape of these functions without changing cumulative demand. For this reason I will impose either a scale or shift normalization. This leads us to the following location and scale identification assumption.

Assumption 2.3. *(i) Assume that contract parameters and demand parameters are either additively or multiplicatively separable in time in the form* $\psi(\cdot, t) = \varphi_1(\cdot) + \varphi_2(t)$ *or* $\varphi_1(\cdot)\,\varphi_2(t)$*, where* $\varphi_1(\cdot)$ *does not depend on time, while* $\varphi_2(0) = 0$*. In case of multiplicative separability an additional scale assumption is* $\varphi_2(1) = 1$*.*

(ii) Assume that $\gamma(t, 0) = 0$ *and* $\gamma(t, 1) = 1$ *so that only the shape of the term defining the heteroscedasticity of the tastes is defined but not its scale.*

I consider here two possible kinds of data. The first kind are the panel data where manager and firm can be observed for several periods. The second kind of data contains i.i.d. observations of firms and managers over time. Identification arguments will be different for these two kinds of data. While the first kind of data allows us to identify the demand and parameters of the contract without imposing the condition of the contract optimality (2.11), the second kind of data requires this condition. The identification results can be summarized in the following theorem.

Theorem 2.1. *Suppose that assumptions 1-3 are satisfied. Then the following con-*

ditions identify parameters of the model for two possible kinds of data:

- *If the data have panel structure where both dimensions asymptotically approach to infinity, then drift and diffusion terms in demand $\psi(\cdot)$, $\zeta(\cdot)$, diffusion term of tastes $\gamma(\cdot)$, contact parameters $\alpha(\cdot)$ and $\beta(\cdot)$, cost function $c(\cdot)$ as well as risk aversion of the manager R_m are identified.*

- *If the data have panel structure then risk aversion of the firm owner R_o as well as fraction of profit in sales δ can only be identified if the contract is optimal, i.e. its components solve (2.11).*

- *If the data have i.i.d. structure, then the components of the model $\psi(\cdot)$, $\zeta(\cdot)$, $\gamma(\cdot)$, $\alpha(\cdot)$, $\beta(\cdot)$, $c(\cdot)$, R_m, R_o, δ are identified only if the contract is optimal, i.e. its components solve (2.11).*

Note that this result basically states that for sufficiently rich data one can identify the parameters of the model without using the conditions of contract optimality. In these circumstances, the problem of the owner adds extra constraints to the model. It is possible, therefore, both to test the validity of these constraints (and, thus, check the optimality of existing managerial contracts), and to build an overidentified model if these constraints appear to be valid.

The Brownian motion assumption restricts the dynamics of tasets and demand process to standard functionals of Brownian motion. This is an important assumption for model identification which simplifies the structure of optimal manager's response. A non-standard process for the demand and tastes dynamics is more complex. This is an area of future research.

The optimal contract mode described by assumptions 1-3 is similar to the class of non-linear models with non-separable errors that have been studied in recent literature. The role of the error term in my model is played by the taste parameter. Identification and estimation of general nonlinear non-separable models were described (among others) in [Mat03] and [Che03]. Identification of my model is facilitated by the panel data structure of my dataset and the assumption that shocks in tastes and demand are uncorrelated. As a result, I can apply an appropriate technique applicable for non-separable models and estimate the parameters of the model.

2.3.2 Estimation

I use a conditional simulation strategy to estimate my model. The idea of estimation is to compare the joint distribution of the observed state variables with the joint distribution of the simulated state variables: salaries and bonuses of managers and firms' sales. Let us first elaborate the mechanism for recovering these joint distributions and then discuss the appropriate methods for their comparison.

The estimation proceeds in three steps. In the first step I flexibly estimate the joint distribution of salaries and bonuses of manager and sales of firms. This estimation is performed conditional on observable covariates. In the second step I simulate from random shocks and compute optimal behavior of the manager conditional on the observable covariates. The optimal strategy of the manager can be computed from the equation (2.9) describing the law of motion of manager's value function. Given the evolution of the taste shocks and demand shocks, I can generate the path of the firm's sales driven by the optimal strategy of the manager and, subsequently, compute bonuses and salaries of managers. This will produce the distribution of sales at each moment for the considered set of firms.

In the third step, I match the distribution of the simulated state variables given the optimal behavior of the manager with the distribution of the observed state variables conditional on the covariates. To match these two distributions, I minimize a distance criterion with respect to parameters describing the structural functions of the model.

Step 1. For manager i D^i_T denotes the total sales in the interval of length T (fiscal year) while W^i and P^i denote the total salary and bonus of the manager. For that period the joint distribution of these variables for the sample of independent observations can be evaluated, for example, using the kernel density estimator:

$$\widehat{f}_{\theta_0}\left(D,\, W,\, P\right) = \frac{1}{n\, h_D\, h_w\, h_p} \sum_{i=1}^{n} \kappa\left(\frac{D - D^i_T}{h_D}\right) \kappa\left(\frac{W - W^i}{h_w}\right) \kappa\left(\frac{P - P^i}{h_p}\right),$$

where $\kappa(\cdot)$ is a kernel function and h_D, h_w and h_p is a set of bandwidth parameters. The following set of assumptions are needed to assure consistency and asymptotic normality of the non-parametric estimates.

A1 The four-dimensional realizations of the stochastic process $\{D^i_\tau,\, W^i_\tau,\, P^i_\tau,\, \theta^i_\tau\}_{\tau\in[0,\,T]}$ are jointly independent across all $i = 1, \ldots, n$ for each $\tau \in [0,\, T]$ and have density functions $\varphi^\tau\left(D_\tau,\, W_\tau,\, P_\tau,\, \theta\right)$ for each $\tau \in [0,\, T]$.

A2 The kernel function $\kappa(\cdot)$ is continuous and bounded and satisfies

- $\int \kappa(z)\, dz = 1$
- $\int |\kappa(z)|\, dz < \infty$
- $\int z^2 \kappa(z) < \infty$

A3 The three-dimensional random variable $\{D^i,\, W^i,\, P^i\}_{i=1}^n$, corresponding to the realization of the stochastic process $\{D^i_\tau,\, W^i_\tau,\, P^i_\tau,\, \theta^i_\tau\}_{\tau\in[0,\,T]}$ at $\tau = T$, has a marginal density which is twice continuously differentiable in its arguments. The derivatives are continuous for $\tau \in [T - \delta,\, T]$ for some $\delta > 0$.

A4 As $n \to \infty$: h_w, h_D, and $h_p \to 0$; $n\, h_D\, h_w\, h_p \to \infty$; $\sqrt{n\, h_D\, h_w\, h_p}\,(h_D\, h_w\, h_p)^2 \to 0$.

Given these assumptions the estimate of the density function is converging pointwise to the true density at the non-parametric rate and is asymptotically normal:

$$\sqrt{n\, h_D\, h_w\, h_p}\left(\widehat{f}(D,\, W,\, P) - f_0(D,\, W,\, P)\right) \xrightarrow{d} N\left(0,\; f_0(D,\, W,\, P)\left\{\int\limits_{-\infty}^{+\infty} \kappa^2(x)\, dx\right\}^3\right)$$

The bias is negligible in the asymptotic distribution because of the undersmoothing assumption in **A.4**. As in standard kernel estimation theory, the consistency result $\widehat{f}\,(D,\, W,\, P) \xrightarrow{p} \int \varphi^T\,(D,\, W,\, P,\, \theta)\; d\theta$ relies on the twice-differentiablity of the marginal joint density. Under the assumption that the bandwidth approaches zero as a function of n, the variance can be found by taking the expectation of the square of the individual term in the kernel density estimate.

Step 2. Conditional on the contract characteristics and the for each value of the structural parameter the distribution of observable state variables is obtained using the following sampling algorithm. Given a partition of time $\{t_j\}_{j=1}^{N_\tau}$ where $t_1 = 0$ and $t_{N_\tau} = T$, simulation increments of the Brownian motion as $\Delta_\tau B_t^D \sim \sqrt{\tau}\epsilon_1$ and $\Delta_\tau B_t^\theta \sim \sqrt{\tau}\epsilon_2$ where ϵ_1 and ϵ_2, are independent standard normal random variables in each elements of partition. These draws are the used to recover the tastes and demand dynamics. While the Brownian innovations in the demand and tastes are independent, they are correlated due to function dependence. The paths of state variables are then computed conditioning on the observable covariates.

A naive simulation for the increments of Brownian motion-driven variable leads to the simulation error of order $\sqrt{\tau}$. However, in order to reduce simulation error I use a second-order Taylor-Itô approximation to improve simulation precision. Higher-order

Taylor-Itô expansions are useful to obtain higher-order precision for solutions of partial differential equations, see for instance [KP92]. For my purposes it is sufficient to consider a second-order approximation, leading to an error of order τ. The dynamics of tastes can be then represented by recursion:

$$\theta_{t_j} = \theta_{t_{j-1}} + \gamma(t_j, \theta_{t_{j-1}})\Delta_\tau B^\theta_{t_j} + \frac{1}{2}\gamma^\eta(t_j, \theta_{t_{j-1}})\frac{\partial \gamma^\eta}{\partial \theta}\left[\left(\Delta_\tau B^\theta_{t_j}\right)^2 - \tau\right]. \tag{2.12}$$

Using the tastes dynamics from equation (2.12), observed level of consumers' demand D_{t_j} at instant t_j given the tastes θ_{T_j} optimal brand characteristic x_{t_j} is obtained by numerically solving equation (2.9). Given the numerically computed brand characteristic, I simulate the sales at time t_{j+1} as:

$$\begin{aligned} D_{t_{j+1}} = D_{t_j} + \psi^\eta(t,\, D_{t_j},\, x_{t_j} - \theta_{t_j})\,\tau + \zeta^\eta(t,\, D_{t_j},\, x_{t_j} - \theta_{t_j})\,\Delta_\tau B^D_{t_j} \\ + \tfrac{1}{2}\zeta^\eta \tfrac{\partial \zeta^\eta}{\partial D}\left[\left(\Delta_\tau B^D_{t_j}\right)^2 - \tau\right]. \end{aligned} \tag{2.13}$$

In equations (2.12) and (2.13) η specifies the parametrization of the functions of the model. The estimation procedure runs recursively up to instant T to obtain the simulated demand evolution and the total sales during the fiscal year $D^{(s)}_T$. Given the contract characteristics and the demand process, the simulated components of manager's compensation conditional on the observed covariates can be obtained as:

$$W^{(s)}_T = \sum_j \alpha(t_j,\, D_{t_j})\,\tau,$$

and

$$P^{(s)}_T = \sum_j \beta(t_j,\, D_{t_j})\left(D_{t_j} - D_{t_{j-1}}\right).$$

As the structural functions determining the evolution of the tastes and the demand process depend on the covariates, each simulated observation will correspond to one actual observation. The density of the simulated total sales of the firm as well as the components of the manager's contract can be again estimated using the kernel estimator:

$$\widehat{f}^{(s)}\left(D,\,W,\,P\right)=\frac{1}{n\,h_d\,h_w\,h_p}\sum_{j=1}^{n}\kappa\left(\frac{D-D_{jT}^{(s)}}{h_s}\right)\,\kappa\left(\frac{W-W_j^{(s)}}{h_w}\right)\,\kappa\left(\frac{P-P_j^{(s)}}{h_p}\right),$$

where $\widehat{f}^{(s)}\left(D,\,W,\,P\right)$ reflects that it is computed from the simulated observations given the structural parameters η. In this expression I omit for brevity the component of the density reflecting conditioning on the observable covariates. The density for the structural parameter η is obtained as:

$$\hat{f}_{\eta}\left(D,\,W,\,P\right)=\frac{1}{S}\sum_{s=1}^{S}\widehat{f}^{(s)}\left(D,\,W,\,P\right),$$

where S is the number of simulations per observation. The problem is now to compare the simulated distribution function $\hat{f}_{\eta}(\cdot)$ and the observed distribution function $\hat{f}_{\eta_0}(\cdot)$. I then estimate the model by finding the parameter η that minimizes the Hellinger distance between the observed and simulated distributions. Chapter 1 uses the Kullback-Leibler divergence for estimation in similar circumstances. In this paper, I use those results and extend them to the estimator described here.

The proof of asymptotic normality of the structural estimates is provided in Appendix B. The proof starts with the analysis of simulation error connected with the approximation of the stochastic integrals by their discretized counterparts. In Appendix B I show that if the step of discretization for simulation of the Brownian motion is decreasing fast enough, then the simulation error will be negligible. The methodology for simulations in equations (2.12) and (2.13) is essential and a similar order of approximation cannot be achieved using a naive sampling scheme. The results in Appendix B show that under certain conditions the discretization error due to simulation is negligible and the only remaining source of error in the density estimation is the sampling noise. This means that the only source of error associated with the evaluation of the density of the joint distribution of sales, wage and bonus of the manager is a standard estimation error encountered in nonparametric density estimation. The estimated distribution will converge to the true distribution at a non-parametric rate. The asymptotic distribution for the simulated data then provides an accurate estimate for the actual distribution of state variables.

2.4 Analysis of incentives in the apparel retail industry

In this section I exemplify the modelling strategy that I develop in the previous sections to estimate structural parameters of executive contracts in the apparel retail industry. I then use the obtained estimates to analyze the effects of the introduction of the Sarbanes and Oxley act.

2.4.1 Data

I analyze the composition of compensation for the managers in the apparel retail industry. The data that I am going to use for analysis are collected from ExecuComp, a part of Standard & Poors' COMPUSTAT dataset, which contains detailed information about the structure of managerial compensations as well as the annual data on firms and compensation of executives. I only use the set of publicly traded companies in the ExecuComp dataset for three reasons. First, it is likely that in the publicly traded companies there is more control of investors over the managerial performance.

Second, the ExecuComp data are particularly interesting because it contains information that allows me to analyze the change in the structure of managerial contracts in response to the introduction of the Sarbanes-Oxley act. Lastly, the publicly traded apparel retail companies are more likely to adopt innovative production and distribution technologies. In addition, vertically vertically integrated production structure of these companies makes it easier to swiftly change the production structure in response to demand shocks.

The ExecuComp data for the apparel retail industry covers 26 apparel retail industry firms. For each firm it includes the compensation data for 7 executive managers per firm on average[5]. My sample includes the top 10 firms in the industry with the largest revenues. In addition to compensations, the dataset includes rich personal data for the managers as well as for the company. The subset of data used in this paper covers the time period from the year 1992 to the year 2005, which spans 12 years before the introduction of the Sarbanes-Oxley act and 2 years after the law became effective.

The company information includes a detail information for company sales, assets and other characteristics. The individual characteristics for manager contain standard demographic variables. The exact manager data available are: date when became CEO, date when joined the company, most recent title, rank by salary and bonus, gender, age, name. These characteristics are displayed in Table B.1[6].

Table B.1 indicates that the average bonus is a large fraction of the base salary,

---

[5] I assume that there no unobserved heterogeneity in the sample, given the observed covariates, that creates no ex ante asymmetry in the data. Conditional on the executive's rank in the firm, it is reasonable to assume that managers solve similar tasks in firms with similar observable characteristics.

[6] The tables in this section are presented in Appendix B

indicating that the role of performance - based payment is significant in the total compensation of executive managers. Most demographic variables are available only for subset of the entire sample[7]. An exception is gender for which I control in the estimation procedure. The gender variable suggests that most executives are males. Females constitute only 21% of the total management. Moreover, the youngest observed age for the manager in the dataset is 42 and the oldest observed age is 78 with an average observed age of 54 years.

Compensation data contain a wide range of components of the incentive contracts. Specifically, the data include the following items: salary, bonus, other annual compensations, option grants, amount of exercised options, restricted stock holdings (number of shares and value), options granted (number of shares and Black-Scholes value), long term incentive payout, and salary change per year. These variables provide the information about the available options and their prices, payments for participation at the executive boards, total cash paid to the managers and a list of other incentive payments such as available stock shares of the company. Summary statistics of the data are displayed in Table B.2.

Table B.2 shows that a significant amount of wealth is concentrated in different forms of options. In addition, there is a large proportion of unexercisable options as compared to exercisable options. The financial performance - based incentives for the managers provided by the option compensation are designed to maximize the performance of the market value of the firm.

[7]Some managers' gender can be inferred from their names and publicly available information source about the industry

2.4.2 Structural estimation

In this section I report the structural estimation results obtained from applying the estimation algorithm developed in the pervious section to the data from the apparel retail industry. Observed heterogeneity in the apparel retail industry is controlled for by data about capital and asset structure of firms which includes the data on assets, market value of shares, profit, revenue from sales. The estimation procedure requires non-parametric estimation of the distributions of observed and simulated sales, bonuses and wages of managers, conditional on the covariates capturing cross-firm heterogeneity. The size of my dataset is not sufficiently large to make precise inference about joint densities of multiple variables. Therefore, I use a linear index $\alpha' Z$ of conditioning variables to capture the dependence of the distribution of state variables from the observed heterogeneity: $f\left(D, W, P \middle| \alpha' Z\right)$. The coefficient in the single index is estimated simultaneously with the structural parameters of the model.

The steps of the estimation procedure described in Section 3 are implemented as follows. For each of the data points, I simulated a pair of Brownian motions with the discretization step $\tau = \frac{T}{100}$. This pair of Brownian motions was fixed for each data point across the parameter search steps. For an initial "guess" of the parameters I evaluated the structural functions of the model: the linear index $\alpha' Z$ in conditional distributions of observable state variables, the drift and the diffusion components of the managerial contract, the drift and the diffusion part of the demand, the drift of the consumers' tastes, and the risk aversion parameter of the managers. Heterogeneous demand and tastes shocks across firms both capture the unobserved heterogeneity in the model. Additional heterogeneity in the demand structure will require additional assumptions about the distribution of random shocks. Introducing an additional unobserved heterogeneity parameter will further increase the computation cost of the model. In the current stage, the model is characterized by 37 parameters.

For each guess of parameter values and for each data point, the optimal manager's strategy is computed by numerically solving the partial differential equation on a square grid. The numerical partial differential equation solver uses the implicit finite difference algorithm, where the parabolic differential equation is transformed into a system of difference equations solvable in a linear number of steps. Equating the expected payoff of the manager at the end of the contract period to the final realized payoff at that time provides a boundary condition for the manager's problem.

Given the solution of the manager's problem, I obtained the optimal branding strategy for each data point. Using the pairs of Brownian motions simulated for each data point together with the optimal branding strategy I recovered the market demand, consumer's tastes, and compensations of the managers at each value of the structural parameters. The joint distribution of the simulated data is then evaluated using a multiplicative kernel density smoother. I obtain the structural estimates by minimizing with respect to parameter values the Hellinger distance between the distribution of the simulated data and the distribution of the actual data. Optimization is carried out using the MCMC, where I use the mean of 100,000 Monte Carlo draws draws for estimation and drop 10,000 of the initial draws.

I demonstrate in Section 3 that the optimal contract model is non-parametrically identified given a large enough dataset. In finite samles a flexible estimation strategy faces the bias-variance tradeoff. In particular, for this model it is computationally intensive to estimate because one needs to compute the optimal behavior of the manager. Thus, the model needs to be flexible but parsimonious enough to produce estimates in a reasonable time. In equations (2.14) - (2.19) I present specifications for the structural functions of the model which I use in the estimation procedure.

The drift of demand is specified as a linear function of time, demand and the distance between the tastes and the brand parameter of the firm:

$$\psi\left(t, D, \theta - x\right) = a_0^{\psi} + a_1^{\psi} t + a_2^{\psi} D + a_3^{\psi} \left|\theta - x\right|. \tag{2.14}$$

The diffusion term is defined similarly to the drift term. This specification implies that the variance of demand shocks can both vary over time and can depend on the actions of the manager:

$$\zeta\left(t, D, \theta - x\right) = a_0^{\zeta} + a_1^{\zeta} t + a_2^{\zeta} D + a_3^{\zeta} \left|\theta - x\right|. \tag{2.15}$$

The dynamics of consumer's tastes is determined by the diffusion term which is specified as a function of time and a current value of the tastes parameter. To capture possible non-linear dependence of variance in tastes from the tastes parameter over time, I include the interaction term:

$$\gamma\left(t, \theta\right) = a_0^{\gamma} + a_1^{\zeta} t + a_2^{\gamma} \theta + a_3^{\gamma} t\theta. \tag{2.16}$$

The shape of the cost function reflects a particular feature of the model considered in this paper: the cost of matching the brand parameter of the firm with the consumers' tastes increases as the brand parameter approaches the tastes. In addition, the cost function is assumed to be symmetric in the distance between the brand parameter and consumers' tastes. Therefore, I specified the cost function as an inverse quadratic polynomial of the distance between the tastes and the brand parameter:

$$c(\theta - x) = \left(a_0^c + a_1^c |\theta - x| + a_2^c [\theta - x]^2\right)^{-1}. \tag{2.17}$$

Finally, parameters of the contract are σ for fixed salary, and functions of demand and time for bonus. Both functions determining the bonus of the manager contain linear terms with time and price and an interaction term, reflecting that the dependence of bonus from the level of sales can change non-linearly over time. This produces:

$$\alpha(t, D) = a_0^\alpha + a_1^\alpha t + a_2^\alpha D + a_3^\alpha tD, \tag{2.18}$$

for compensation rewarding high level of sales and

$$\beta(t, D) = a_0^\beta + a_1^\beta t + a_2^\beta D + a_3^\beta tD, \tag{2.19}$$

for compensation rewarding sales growth. Since it is conventional to assume risk-neutrality of the representative sharedholder, I chose to fix her risk-aversion parameter to be equal to zero. The risk-aversion parameter of the manager was estimated. Parameter estimates obtained from an MCMC estimation procedure are presented in Tables B.3 and B.4.

In estimation I recovered parameters of structural functions (2.14) - (2.19). The estimates of demand components show that demand is increasing within the period as indicated by the parameter a_0^ψ. If the brand parameter is further from the tastes of consumers, the growth rate is slower, as indicated by a_3^ψ. Coefficient a_3^ψ indicates that in order to achieve a high expected level of consumer demand at the end of the period the manager needs to match the consumer's tastes during the entire period.

From parameter a_1^{ζ} one can see that the variance of demand is decreasing during the period while the further the brand parameter is from consumers' tastes, the higher is the volatility of demand over time (indicated by positive value of parameter a_3^{ζ}). In total, demand parameters indicate that matching consumer tastes by the brand parameter of the firm the manager not only maintains high total sales of the firm but also reduces the volatility of sales within the period.

Parameters for the dynamics of tastes show a significant degree of heteroskedastisity in its behavior. In particular, parameter a_1^{γ} indicates decreasing variance of taste innovations over time, but positive a_2^{γ} leads to an increasing variance of taste innovations when the taste parameter is far from zero. The latter coefficient indicates that the stationary distribution of tastes has a substantial probability mass at the tails. As a result, prediction of future tastes becomes a complex statistical problem explaining high cost of tastes monitoring over time.

The cost function shows that it is increasingly costly to narrow the gap between the brand parameter and consumer tastes. Due to symmetric specification of the cost function, the maximum cost is achieved when tastes of consumers are perfectly matched by the brand parameter of the firm. The elasticity of cost with respect to the distance between tastes and the brand parameter linearly decreases as the distance increases.

Lastly, parameters for the characteristics of managerial contracts show that a significant response to the changes in the state variables is demonstrated in long-term compensation, which makes the manager put more emphasis on increasing the growth rate of sales at the end of the period (as indicated by a_1^{β}).

Additional estimated parameters include the wage and the risk aversion parameters. For estimation purposes I normalized observed variables by equating the highest observed values of state variables to 100. Therefore, using the ratio of average total compensation to the maximal total compensation, the average relative risk aversion of the manager is equal to 7.4 which is close to existing estimates of risk aversions of stock managers.

The next set of results represent the parameters in the linear index as well as the normalizing constant for the contract and the risk aversion parameter of the manager. Estimates of the parameters in the single index show that the distributions of observable characteristics of the model are most affected by firms' market values, amounts of option grants awarded to the managers, shares owned by the manager and gender of the manager. These results suggest that structure of manager's compensation is most significantly differ by the size of the firm, the degree of firm ownership by the manager, and potential future ownership in terms of option grants.

2.4.3 Analysis of consequences of the Sarbanes-Oxley act

Overview of changes

The Sarbanes-Oxley act (SOX) was passed by the US senate in July of 2002 as a response to a wave of corporate scandals involving large companies such as Enron and WorldCom. Before the SOX, some disclosure requirements on executive managers were set by the federal regulations while the rest were set by the state legislation. The provisions in the SOX set rigid restrictions on the actions of executive managers and the relation between corporations and auditors, and audit firms. Specifically, the SOX requires that the corporate audit must be conducted by an independent committee. The corporation is prohibited from buying non-audit services from the auditors. Corporate loans to the executive managers are prohibited. It is also required that the executive managers certify the corporate financial statement. In addition, the SOX requires that executive managers return the incentive-based part of the compensation from stock sales if the corporation reinstates earnings. In general, provisions of the Act impose more formal legal restrictions on the employment contracts of the executive managers.

The SOX has received a significant amount of criticism in the theoretical literature. [Rom05] argues that the SOX was implemented during the corporate crisis and thus was not a result of multilateral objective consideration. For this reason, [Rom05] suggests that the SOX has set several unnecessary restrictions which do not prove to be effective in controlling malpractice by executive managers. Rather, it seems mostly to raise the compliance cost for the corporation. More importantly, [Rom05] argues that the SOX has placed a relatively higher compliance cost burden on smaller corporations. Another concern reflected in [CDL04] is that, due to the excessive restrictions in the SOX, the legislation might divert the decision of the managers from value-maximization. They can choose less risky projects in order to reduce personal risk exposure (deadweight loss from sub-optimal decisions). It may even lead corporations to switch to less public forms of ownership. The authors argue that, in response to increased managerial liability, the contracts of the managers should have a proportionally larger fixed component of the manager's compensation. [Rib02] also criticizes the SOX, suggesting that the Act adds restrictions which do not set the correct incentives for the managers and cannot be considered an efficient instrument of prevention for corporate fraud. The general opinion in the literature is that the deadweight loss connected with the act's implementation and enforcement is very large (the estimates of losses by stock market investors in [Zha05] amount to $1.4 trillion) while the benefits in terms of the risk reduction are limited.

The penalty imposed by the SOX leads managers to behave more cautiously in implementing policies. In the data for the years after the SOX was implemented I find evidence consistent with increased their effective risk aversions. In this section using the ExecuComp dataset for the years 2004 and 2005 when the SOX became effective I find the change in the risk aversion consistent with the observed change in the industry. In addition, I will use the dynamic model to predict the long-term effects of the SOX on the industry.

Effect of the SOX on the compensation structure

The Sarbanes-Oxley Act imposes rigid restrictions on managerial performance resulting in substantial loss of compensation in case of undesirable economic outcome. Institutional changes associated with the introduction of the SOX can be modelled by considering an increase in the "effective" risk aversion amongst both managers and the investors. From the analysis of comparative dynamics of my model, I conclude that an increase in the manager's risk aversion should lead to an improved performance and, subsequently, to an increase in her compensation. For the Compustat data the effect of the SOX on managerial compensations can be illustrated by a simple linear regression. In Table B.5 I report the results of 2SLS regressions in which the dependent variables are the normalized sales of the firm (sales minus the average sales on the market normalized by the standard deviation of sales on the market in a specific year). Table B.5 contains three sets of results. The first two sets represent the relationship between salaries and the bonuses of managers from the firm sales. The last set of results contains the estimates for the relation between the sales of the firm and firm's characteristics. In this last model I am concerned with possible endogeneity of such characteristics as the firm's value and I use relative variables such as capital structure of the firm as instruments. The data sample includes the firms doing retail sales business. As covariates I use the firm's market value, total assets, Tobin's q, and change in the assets within the year., with dummy variables for the firms specializing in the apparel retail. I also use a dummy variable for the years 2004 and 2005 when the SOX became effective. Table B.5 demonstrates my estimation results for the system describing the salaries and bonuses of managers across firms in conjunction with sales of the firm. It shows that bonuses and salaries tend to grow with the normalized sales of the firm. More importantly, the introduction of the SOX led to an increase of $27,624 in the base salaries and a $43,494 raise in average bonuses of executive managers. An important economic question is whether compensation has been increased solely as a behavioral response to changing institutional environment or as a response of the structure of the industry to the introduction of the SOX. This question can be explicitly answered by my model. Specifically, the model can be used to predict the changes in the managerial compensation in response to changes in the manager's risk aversion. It is possible to find the change in the risk aversion that will generate the observed change in compensation. The new value of risk aversion will indicate the extent of managerial response which can explain the observed increase in the compensation.

Predictions for industry dynamics

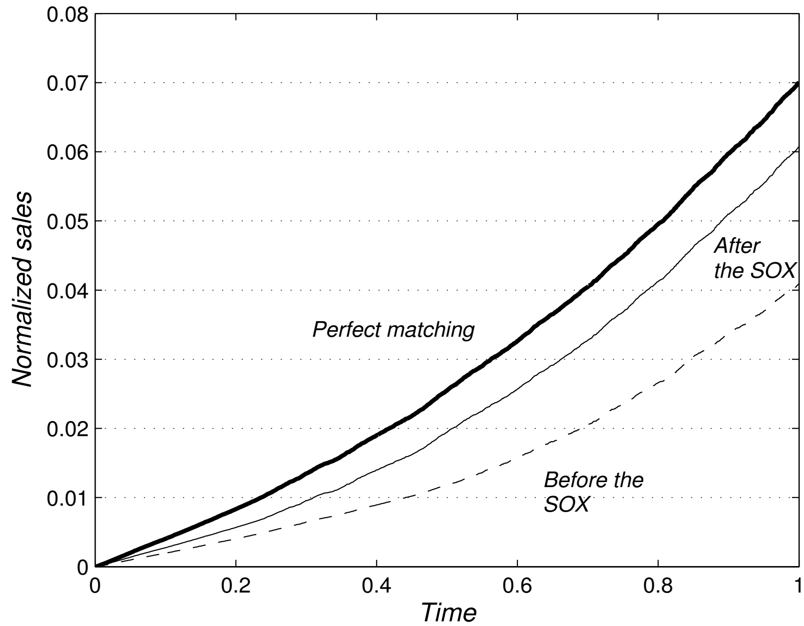
In the previous discussion I recover the effect of changes in the institutional environment on the "effective" risk aversion of executive managers. My continuous-time model of a manager in the apparel retail industry allows me to predict the effect of such an increase on the dynamic path of aggregate sales of the firm as well as the compensation package of the manager. Based on the simulated paths of the sales and the tastes parameter I find a change in the risk aversion which matches the observed change in the compensation to the change in the compensation predicted by my structural model. I use mean-squared deviation of the simulated compensation from the average observed compensation as a criterion for search of the risk aversion and minimize this objective using a standard technique.

In simulations I assume that the risk aversion of the representative shareholder remains the same. The value of the absolute risk aversion of the manager that generates the observed change in the average managerial compensation is 4.02 which is 4 times bigger that the absolute risk aversion estimated on the data before the SOX. This is a very significant change in the risk aversion which was obtained holding the contract structure constant and it might have been a result of structural changes in the industry. However, observations from my data indicate that the profit of firms in the industry increases after the introduction of the SOX. Given that the period that I am considering is short, there was little incentive for the owners of the firm to reconsider the contract structure. My experiment in this case provides a lower bound for the effect of the introduction of the SOX on the long-term sales.

Next I simulate the effect of the estimated change in the absolute risk aversion on the dynamics of the sales of the firm. Simulations for the firm's sales dynamics are obtained directly from simulating (2.12) and (2.13) forward and then averaging over simulated paths. I used three cases to quantify the effect of Sarbanes and Oxley act on the sales of the firm. The first case, which was used as a reference point, is the case when the manager perfectly matches consumer tastes. On average, this case should yield the highest sales. The second case is the case of a risk-averse manager with the risk-aversion parameter obtained from the structural estimation. The third case reflects an increase in the effective risk aversion of the manager due to the introduction of the SOX. Figure 2.1 demonstrates the dynamics of average cumulative sales corresponding to the three cases. One can see that the case of perfect matching of consumer's tastes leads to the highest sales.

The case with the estimated risk aversion of the manager before the SOX leads

Figure 2.1: Predicted cumulative sales of the firm



to the smallest sales. One can also see that the effective increase in the risk aversion of the manager leads to an increase in sales, such that cumulative sales at the end of the period approach to the sales with perfect matching of tastes.

To summarize the simulation exercise, it is clear that the change in the effective risk aversion predicted from the structural model in response to the introduction of the Sarbanes and Oxley act is extremely large. In the model where contract structure remains the same after the legislative change the only variable affecting the amount of compensation is the effort of the manager. It is clear that the changes in the effort that would generate large increases in the compensation have to be induced by a large increase in the risk aversion. An alternative effect which can explain the observed compensation increase is the change in the contract structure induced by response in corporate structure and the structure of ownership of the firm. In general one could expect a combined effect as a result of changing both the effort of the manager and the structure of contracts. The period of observation of the post-SOX effect is relatively short, suggesting that the structure of the contract might not have completely reflected the institutional change. My predictions provide an explanation for one component of the overall response.

2.5 Conclusion

Modeling the dynamic principal-agent problem allows one to analyze the effect of incentives in the agent's contract on her long-run and short-run performance. In this paper I develop a methodology for implementing and structurally estimating a general continuous-time principal-agent model in the presence of the unobserved state variables. The non-linear response of the manager that I observe in my model extends the existing literature and it is more suitable for empirical implementation.

In this paper I provide conditions for identification of the continuous-time model from the data and I develop a methodology for empirical estimation of this model. Estimation methodology is based on matching the observed conditional distributions of state variables with the simulated distribution, for a particular set of structural parameters. Based on these distributions I use the Hellinger distance as a criterion for parameter estimation. I minimize the distance using a Markov Chain Monte Carlo technique.

I estimate the model for managerial contracts in the apparel retail industry. During the estimation procedure I recover the parameter of consumers' demand and characteristic of consumers' tastes. I also estimate the parameters of managerial contracts. I use the obtained estimates to analyze the effect of introduction of the Sarbanes and Oxley act. I find that the post-SOX compensations demonstrate a significant increase after the act. I assume that additional legislative restrictions set by the SOX are increasing the effective risk aversion of the manager. Then I search for an increase in the risk aversion which could generate the observed increase in managerial compensation. Based on the estimated increase in the effective risk aversion I evaluate the effect on expected sales in the industry and I find that after the introduction of the SOX sales of a typical firm should increase. This analysis can be applied to other industries to perform welfare analysis and analysis of the output dynamics.

Chapter 3
Estimation of Continuous-Time Models of Interactions

3.1 Introduction

Resent research shows both practical and theoretical importance of models of dynamic games for industry analysis. One widely used approach to modelling dynamic games follows [EP95] where authors use Markov perfection to describe a computationally tractable dynamic game model. However, there are two min obstacles for implementation of this modelling methodology to a wide range of industries. First, computation of optimal strategies becomes problematic when the number of player is large. Second, it requires a stationary industry environment to assure that the law of motion of state variables is Markovian. One attractive approach to overcome the first problem is presented in [WBVR05]. In [WBVR05] authors suggest substituting particular characteristics of rivals in a very large game by some aggregate features of their actions. This approach, however, does not address the second problem and does not take into account potential non-stationarity in the industry. Moreover, it requires a certain degree of homogeneity of players in the game.

In my paper, I am addressing the question of analyzing a dynamic game with a large number of players in a non-stationary environment. I use an approach similar to that in [WBVR05] and assume that players in the game only respond to aggregate characteristic of rival actions. However, one main distinction of my approach is that I assume that players have different information about the process driving this aggregate state variable. In this case players have an incentive to acquire information about the evolution of the state variable during the game to increase the expected payoff at the end of the game. This assumption makes the game non-stationary because the information sets of players at the end of the game are different from the sets in the beginning of the game. Moreover, in this setting I can analyze learning which in this context is the information acquisition about the dynamics of the state variable.

This structure of the model facilitates its implementation in continuous time. Convenience of considering a dynamic game in continuous time has been noted in [DJ05], however, that paper also analyzes estimation of continuous-time models in the stationary case. In my paper I suggest a natural generalization of this approach based on the transformation of the Bellman equation into the partial differential equation describing the law of motion of the value function. The solution of this equation will produce both the optimal strategy and the value function without regard to the stationarity of the state variables (given that the equation for the dynamics of the state variables is correctly specified). An additional advantage of this approach is that it is not based on any iterative schemes and only relies on the linear methods for integration of partial differential equations. This implies that the number of steps required to find the solution is pre-determined. Such approach leads to a natural simulation-based estimation method in the flavor of indirect inference of [GMR93] and [GT02], based on the comparison between the optimal strategic response to the movement of the state variable and the observed response. Such combination of the continuous-time modeling and the estimation procedure creates a methodology which can be used to analyze models of industrial dynamics in non-stationary environment.

In this paper I apply the developed methodology to analyze Internet auctions. On Internet auction websites, multiple auctions for the same good occur simultaneously. For this reason, bidders choose between the menu of auctions and their entry decisions depend on current prices in the auctions. On the other hand, bidders cannot observe their rivals and have to form predictions about the number and skills of potential rivals in a particular auction. In this case a continuous-time model provides a natural framework for analysis of strategic behavior of bidders without complications inherent to discrete-time models. The state variable in the model is represented by the price in the auction, and the strategy of the bidder is sequence of bids over time, which she chooses to influence the price optimally and maximize her expected surplus from winning the auction.

Contributions of the paper include the development of the framework which allows one to estimate continuous-time partial equilibrium models. In particular, I develop the asymptotic theory for the constructed estimator and prove the main convergence results. In addition, I demonstrate sufficient conditions for existence and uniqueness of equilibrium in Poisson-driven models without direct spillovers between the agents.

The model of Internet auctions which I use for application appears to be non-parametrically identified as shown in Chapter 1. In this paper I give examples of two simple situations where identification is clear. In particular, if there are discrete types of bidders, then it will be possible to predict bidder types from the data by observing their response to price movement. The unobserved visibility of the auction can be captured ex-post by the number of bidders entered into the auction.

The estimator developed in this paper is based on equilibrium considerations. Specifically, the idea of the estimator is that in equilibrium the optimal response of the model (combined optimal response of the participants of the dynamic game) to the movement of the state variable should replicate this movement. In case of the second-price auctions in the Internet, this implies that, if we observe the price movement, and we can simulate the entry of bidders, then we can compute optimal bidding strategies and the second-highest bid. In equilibrium, the distribution of this second-highest bid should coincide with the distribution of the price (equal to the second highest bid) observed from the data. In practice, I suggest to measure the "closeness" of these two distributions by the Kullback-Leibler Information criterion (also known as the K-L divergence). The estimator is based on minimizing the K-L divergence between the distribution of the data and the distribution of the simulated response over structural parameters.

Under certain regularity conditions it is possible to provide the asymptotic normality of the estimates. The convergence properties of the estimator are based on the regularity results derived for estimation of parameters of regressions for count and duration data (Cox regression). A defining difference the estimator in this paper and the case of Cox regression is that my estimator is based on the minimization of a non-parametrically defined objective function. Therefore, the estimator has a non-parametric convergence rate.

I suggest an estimation method based on the MCMC for the quasi-posterior defined by the K-L divergence, and derive the relation between the asymptotic results and the estimation outcome of the MCMC procedure. It appears that the MCMC gives an unbiased estimator, but it underestimates its variance. Computations show that the actual variance of the estimator can be obtained by doubling the variance estimates from the MCMC procedure.

The performance of the estimator is evaluated in a number of Monte-Carlo experiments. The experiments confirm the convergence of the estimator and show good coverage of the asymptotic confidence intervals.

The continuous-time framework provided in this paper allows one to make welfare comparisons between different setups of continuous-time games. In the example of continuous-time auctions we can compare the welfare of the bidders and sellers for different auction formats. Two dominant formats in the Internet are the second-price format with a fixed ending time on eBay ("eBay format") and the flexible-ending format used mainly by the Amazon.com website ("Amazon format"). In the auction with the Amazon format the ending time of the auction is driven by the bidder activity: the auction continues until new bids stop arriving into the auction. I incorporate this feature into the model of a continuous-time auction considered in this paper. Then welfare comparisons between two different auction formats are obtained on the basis of structural estimates of the auction features for eBay auctions of music CDs.

The structure of the paper is the following. Section 2 provides a general setup of continuous-time modeling and gives a particular example of a continuous-time model of an Internet auction. It also proves the existence and uniqueness of equilibrium in the considered model. Section 3 discusses identification of the continuous-time model and gives examples of identification in the model with different structures for the unobserved heterogeneity. Section 4 derives the asymptotic behavior of the suggested estimator and provides guidance for obtaining the asymptotic standard errors from the MCMC procedure. This section also demonstrates the performance of the estimator in a series of Monte-Carlo experiments. Section 5 demonstrates how the suggested continuous-time framework can be used for welfare and revenue comparisons for the example of eBay and Amazon auction formats. Section 6 concludes.

3.2 Model

3.2.1 Setup

In this paper, I consider a general model of continuous-time strategic interactions in a quasi-competitive environment. In this environment individual participants do not directly observe the identities and actions of their rivals. Rival behavior is summarized by a vector of state variables. I assume that the dynamics of these state variables are described by a continuous-time Markov process, coming from a certain parametric family. This family is assumed to be known to the players. Each player forms a best response to the state variable process. In this way, the equilibrium in this model has a competitive structure: players best-respond to the dynamics of the state variable and their aggregate best responses shape the dynamics of the state variable.

To develop this concept formally, I specify timing of the game, players, structure of actions and payoffs for the players, and then describe the dynamics of the endogenous state variable. I assume that the game starts at time $t = 0$ and ends at time $t = T$, which is known to each player i. The number of rivals is unknown to players. The actions of players $a_{it} \in \mathcal{A}_i$ belong to some bounded continuous or discrete set. All players observe the state variable $x_t \in \mathcal{X}$. The dynamics of this state variable is described by a controlled Markov stochastic process with piecewise-continuous sample paths $Z_t\left(\cdot, x_t, \theta\right)$ such that

$$dx_t = Z_t\left(dt, x_t, \theta\right). \tag{3.1}$$

It is a common knowledge that $Z_t\left(\cdot, \theta\right)$ belongs to a parametric family of stochastic processes $\mathcal{Z} = \left\{Z_t\left(\cdot, \theta\right), \, \theta \in \Theta\right\}$ characterized by parameter $\theta \in \Theta \subset \mathbb{R}^p$. In general, the following discussion is relevant to any $Z_t(\cdot, x_t, \theta)$ driven by a diffusion or a jump-diffusion process. However, for clarity of exposition we will consider the case when $Z_t(dt, x_t, \theta)$ is driven by the increments of the heterogeneous Poisson process:

$$dx_t = h(t, x_t, a_t)\ dJ(t, x_t, \theta), \tag{3.2}$$

where $J(t, x_t, \theta)$ is a Poisson process with the frequency $\lambda(t, x_t, \theta)$. Therefore, the class $\mathcal{Z}$ contains all Poisson processes with frequencies $\lambda(t, x_t, \theta)$.

The value θ_0 characterizing a particular process is not known to players. This parameter is designed to capture a form of unobserved heterogeneity across different games driven by stochastic processes from the set $\mathcal{Z}$. Such a parameter can be interpreted as a characteristic of equilibrium selection or selection of player types into the game. In general, uncertainty of players about θ can also be interpreted as their uncertainty about the "rules" of the game. Heterogeneity of information about the parameter θ would then be a characteristic of players' skills. I assume that players have initial beliefs about θ in the form of normal distributions with means μ_i and standard deviations σ_i. The distribution of beliefs across players is described by distribution functions $F_\mu(\cdot)$ and $F_\sigma(\cdot)$. We assume that the players can learn about θ_0 by observing the behavior of the state variable. I consider second-order Bayesian learning when players only update means and variances of their beliefs (but not the entire distribution) during the game. The vector of state variables for player i at time t is a triple $(x_t,\ \mu_{it},\ \sigma_{it})$. Strategies of players define a mapping between the vector $(x_t,\ \mu_{it},\ \sigma_{it},\ t)$ and actions $a_{it} \in \mathcal{A}_i$ for all $t \in [0,\ T]$. Players maximize their utilities at the end of the game $u_i(x_T, a_{iT})$[1] by choosing optimal actions given the dynamics

[1]It is possible to generalize this model to the case when utilities of players also depend on the

of the state variable (3.2) and transition of beliefs.

The number of players in the game will be described by the relation

$$N_t = \nu\left(x_t,\, \xi_t,\, t\right), \tag{3.3}$$

where ξ_t is a random variable such that $\Pr\left\{Z_t\left(t, x_t, \theta\right), \xi_t\right\} = \Pr\left\{Z_t\left(t, x_t, \theta\right)\right\}$. This formulation allows for cases when the number of players in the game is fixed. On the other hand, if the number of players is endogenous, I assume that both the state variable and the number of players are affected by the same shock.

In this model I consider an equilibrium where the the probability distribution for the generating process $Z\left(\cdot, x_t, \theta\right)$ is exogenous. In particular if the model is generated by a jump-diffusion process, it is assumed that the Brownian motion and the Poisson measure generating jumps are exogenous. Examples of such exogenous processes could be stochastic entry into the game, stochastic demand in a differentiated product market, stochastic productivity, etc. The market-clearing conditions in this model relate the moments of the stochastic process $Z\left(\cdot, x_t, \theta\right)$ with actions of players. Suppose that $E_t(a_{it})$ is the expected value of state variable x_t, $V_t(a_{it})$ is the variance of state variable x_t, and $Q_t\left(a_{it}\right)$ is the quadratic variation of x_t given action of player i at time t. Equilibrium conditions take the form:

$$\varphi_i\left(E_t\left(a_{it}\right),\, V_t\left(a_{it}\right),\, Q_t\left(a_{it}\right),\, a_{-i,t},\, t\right) = 0, \text{ for } i = 1, \ldots, N_t, \tag{3.4}$$

where N_t is the number of players in the game at the instant t. This expression

intermediate values of state variable and actions. In this case the time of entry into the game will determine the type of a player in addition to the values of beliefs.

can take a simpler form if one considers a pure jump Poisson process as in (3.2). In particular, the equilibrium condition can then be written as:

$$\varphi\left(h\left(t, x_t, a_{it}\right), a_{-i,t}\right) = 0, \text{ for } i = 1, \ldots, N_t.$$

I assume that equilibrium conditions do not over-identify the model. In particular, the set of solutions to (3.4) $\mathcal{S} \subset \mathcal{A}_1 \times \ldots \times A_{N_t}$ is not empty and the Jacobi matrix $\nabla_{a_i}\varphi_i\left(E_t\left(a_{it}\right), V_t\left(a_{it}\right), Q_t\left(a_{it}\right), a_{-i,t}, t\right)$ does not have a full rank for almost all t. If system (3.4) does not have solutions, the model does not have any equilibria. I assume that (3.4) has at least one solution to be able to consider non-trivial dynamics in the game.

3.2.2 Solving for the Best Response of Players

Players in the considered game maximize their expected utility at the end of the game given laws of motion for the state variables of the model: x_t, as well as means and variances of their beliefs. The law of motion for x_t is given by a stochastic differential equation (3.2). To close the optimization problem of the player I provide the law of motion for beliefs about the parameter θ_0. The law of motion of players' beliefs about θ_0 can be interpreted as the description of "learning" during the game. For brevity, I abuse notation in this section and omit index i referring the strategy and actions of a particular player.

To simplify the learning problem I use the linear Bayesian learning approach. In general, each innovation in the state variable x_t should change the entire distribution of player's beliefs about θ_0. However, the problem of non-parametric updating of beliefs is frequently intractable. One way of solving this problem is to substitute the problem of Bayesian learning for the entire posterior distribution by the problem of learning about a particular fixed set of moments of posterior distribution. In this case the law of motion for these moments can be represented in the form of a system of linear stochastic differential equations. These equations are typically referred to as a linear filtration equations in the statistical literature. Linear filtration equations for jump-diffusion processes can be found in [LS01]. In this paper I use an example of of a Poisson process and to compute the dynamics of players' beliefs I use the linear filtration method developed in [VS77]. An optimal linear filter (as shown in [LS01]) is infinite-dimensional. I choose to analyze only the dynamics of posterior means and variances of players' beliefs, which reduces infinite-dimensional system to two equations. The following theorem establishes this result formally.

Proposition 3.1. *Let $\mathfrak{F}_{xt}$ be σ-algebras generated by the sample trajectories of the state variable process. Assume that $E\{(\theta_t - \theta_0)^3 | \mathfrak{F}_{xt}\} = 0$ for any $t \in [0,\, T]$. Moreover, assume that the compensated state variable process has an integrable quadratic variation. Then the dynamics of the mean and the variance of player beliefs are given by:*

$$\begin{aligned} d\mu_t &= \tfrac{\partial \lambda(t, x_t, \mu_t)}{\partial \theta} \tfrac{\sigma_t}{\psi(t, x_t, \mu_t)} \left\{ dx_t - \psi(t,\, x_t,\, \mu_t)\, dt \right\}, \\ d\sigma_t &= -\tfrac{1}{\lambda(t, x_t, \mu_t)} \left(\tfrac{\partial \lambda(t, x_t, \mu_t)}{\partial \theta} \right)^2 \sigma_t^2 \, dt. \end{aligned} \tag{3.5}$$

In this expression $\psi\,(t,\, x_t,\, \mu_t) = h\,(t,\, x_t,\, a_t)\, \lambda_\epsilon\,(t,\, x_t,\, \mu_t)$ characterizes the compensator of the state variable process (see [LS01]). The proof of Proposition 3.1 is given in Appendix C.

Changes in the mean of player beliefs regarding parameter θ are connected to changes in the observed movement in the state variable by the first stochastic differential equation in system (3.5). This equation suggests that changes in the mean of player beliefs become larger if the variance of these beliefs grows. If the expected growth rate of state variable x_t increases, changes in player beliefs become relatively smaller. The second equation in system (3.5) describes the dynamics of the variance of beliefs. This equation shows that the variance decreases over time, implying that players' information about parameter θ becomes more precise towards the end of the game[2]. According to equations (3.5), players re-evaluate their strategies after each jump in x_t. If their information about θ was accurate, then this adjustment will be small. However, if prior beliefs are diffuse, then each additional jump in the state variable will affect strategies significantly. Equations (3.5) close the optimization problem of an individual player. I proceed with the formal solution of this problem.

As I stated in the previous section, players maximize their utilities at the end of the game given dynamics of state variable x_t and beliefs about θ identifying the process for x_t in the parametric family $\mathcal{Z}$. This is a stochastic dynamic optimization problem with initial conditions defined by the value of the state variable in the beginning of the game x_0 and initial player beliefs. This problem can be expressed as:

$$\begin{aligned}
&\max_{a_t}\ E_0\{u\,(x_T,\,a_T)\}\\
&dx_t = h\,(x_t,\,t,\,a_t)\ dJ(t,\,x_t,\,\mu_t),\\
&d\mu_t = \tfrac{\partial\lambda(t,\,x_t,\,\mu_t)}{\partial\theta}\tfrac{\sigma_t}{\psi(t,\,x_t,\,\mu_t)}\left\{dx_t - \psi(t,\,x_t,\,\mu_t)\,dt\right\},\\
&d\sigma_t = -\tfrac{1}{\lambda(t,\,x_t,\,\mu_t)}\left(\tfrac{\partial\lambda(t,\,p_t,\,\mu_t)}{\partial\theta}\right)^2\sigma_t^2\,dt.\\
&\theta_t\big|_{t=0} \sim N\left(\mu_0,\,\sigma_0\right),\quad x_t\big|_{t=0} = x_0.
\end{aligned} \tag{3.6}$$

[2]The system (3.5) shows a significant difference between the learning concept considered in this paper and a commonly used model of reinforcement learning. In this paper the object of learning by the player is the optimal playing strategy and convergence of the playing strategy to the optimal playing strategy is justified by the concentration of player beliefs about parameter θ in the vicinity of the true parameter over time. The differences in the speeds of learning between the different players are ruled by the initial differences in the beliefs such that the bidders who are less certain about parameter θ in the beginning of the game will be learning about the optimal playing strategy slower than their more informed rivals.

The first line of (3.6) contains the objective of the player (expected utility at the end of the game), the second line describes the changes in the state variable, observed by the player, the third and the fourth line give the dynamics of player beliefs, and the the last expressions give the initial conditions to the problem.

The solution to this dynamic optimization problem can be found using Itô calculus for Poisson-driven processes. To do so we define the value function of the bidder as:

$$V\left(t, x, \mu, \sigma\right) = E_t \left\{u\left(x_T,\ a_T\right)\right\}.$$

Function $V(\cdot)$ specifies the expected surplus of the player from winning the auction given that, at time $t \in [0, T]$, state variable is equal to x while distribution of player beliefs about θ has mean μ and variance σ. The introduction of this value function allows us to reduce the problem of stochastic optimization (with the initial conditions determined in general by the random variables), to the non-stochastic partial differential equation. The natural boundary condition to this equation is provided by the fact that the value function at the end of the game is equal to the terminal utility (as at the terminal instant all uncertainty has realized). Standard derivations (following, for instance [GS79]) give the the law of motion of the value function of the player:

$$\begin{aligned}
&\tfrac{\partial V(t,x,\mu,\sigma)}{\partial t} + \sup_{a\in\mathcal{A}} \Big[-\tfrac{\sigma}{\lambda}\tfrac{\partial\lambda}{\partial\theta}\tfrac{\partial V}{\partial\mu} - \tfrac{\sigma^2}{\lambda}\left(\tfrac{\partial\lambda}{\partial\theta}\right)^2\tfrac{\partial V}{\partial\sigma} + \\
&+V\left(t,\ x+h\left(t,x,a\right),\ \mu+\tfrac{\partial\lambda}{\partial\theta}\tfrac{\sigma}{\lambda},\ \sigma\right)\lambda\left(x+h\left(t,x,a\right),\ \mu+\tfrac{\partial\lambda}{\partial\theta}\tfrac{\sigma}{\lambda},\ t\right) - \\
&-V(t,\ x,\ \mu,\ \sigma)\,\lambda\left(x,\ \mu,\ t\right)\big] = 0
\end{aligned} \tag{3.7}$$

$$V(T,\ x,\ \mu,\ \sigma) = \sup_{a_T\in\mathcal{A}} u\left(x_T,\ a_T\right)$$

I assume that the space of admissible controls $\mathcal{A}$ is such that $a \in \mathcal{A}$ are continuous and differentiable almost everywhere functions defined on $[0, T] \times [0, \overline{X}] \times \Theta \times \Sigma$. I assume that these functions are bounded by some $\overline{X} < \infty$ and that $h(\cdot, a_t)$ are measurable functions of finite variance with respect to the Poisson measure. The system (3.7) represents the boundary problem for the generalized heat equation for the value function of the player $V(\cdot)$ often classified as a Dirichlet problem.

3.2.3 Equilibrium

I have outlined previously the structure of a quasi-competitive environment of the model. In this environment individual players solve dynamic optimization problems and form responses to observed movement of the state variable. The profile of optimal actions of players is related with the movement of the state variable through generalized market-clearing conditions (3.4). Uncertainty in the model is captured by limited information of players about parameter θ characterizing a particular process generating the state variable x_t from the parametric family of processes $\mathcal{Z}$. I require that beliefs of players across the population are stationary over time. This means that at any given time, means and variances of beliefs of players have the same distribution at any instant of the game. This can be interpreted as the fact that players do not make systematic mistakes in their estimates of parameter θ. In addition I postulate that learning in the model occurs according to equations (3.5). In this way, first, I automatically assure that, for consistent prior beliefs, the mechanism will produce consistent posterior beliefs. Second, this learning mechanism provides refinement of information about θ over time if initial beliefs are not degenerate. This feature in particular allows all players to get perfect information about the process generating the state variable if the duration of the game approaches infinity.

These considerations are summarized in a formal definition of an equilibrium.

Definition 3.1. *A competitive dynamic equilibrium is a trajectory of state variable x_t, the number of players N_t and a collection of means and variances of player beliefs $\{\mu_t^i,\ \sigma_t^i\}_{i=1}^{N_t}$ for players indexed $i = 1, \ldots, N_t$ for instant $t \in [0,\ T]$ such that:*

- *Number of players in the game N_t is determined by (3.3).This number is not observed by the participating players. Both the number of players and the state variable are driven by the same shock.*
- *Initial beliefs of players are $\mu_0^i \sim G_\mu(\cdot)$ and $\sigma_0^i \sim F_\sigma(\cdot)$, and utility functions of players $u_i(\cdot)$ are determined by the index of player*
- *Players solve the dynamic optimization problem (3.6) which generates their best responses which map observed state variable movement and beliefs into actions: $a_{it}^*(x, \mu, \sigma)$ and the paths of beliefs $\{\mu_t^i,\ \sigma_t^i\}_{t\in[0,\,T]}$ with $\mu_t^i|_{t=0} = \mu_0^i$ and $\sigma_t^i|_{t=0} = \sigma_0^i$ for players $i = 1, \ldots, N_t$ for each $t \in [0,\ T]$*
- *For each t optimal actions of players a_{it}^* satisfy generalized market clearing conditions (3.4).*

As there is no strategic interaction between different players in this model, the property of the equilibrium reflects the property of solution to the optimal bidding problem. The fact that the price movement is self-reinforcing makes a competitive dynamic equilibrium a fixed point of the stochastic differential equation (3.7).

The following theorem establishes that the equilibrium exists and is unique.

Theorem 3.1. *Suppose that the stochastic differential equations, representing the dynamics of state variables (i.e. x_t and player beliefs) have continuously differentiable coefficients and square - integrable solutions with probability one. In this case the equilibrium represented by Definition 3.1 exists and is unique.*

Proof:

I will prove this theorem for a particular case of a pure jump process. However, the existence argument remains valid for a general jump-diffusion process for the state variable, while the uniqueness will hold if the process is drift-dominant (the mean square error of shocks over time is smaller than movement along the non-linear trend).

Existence. In Appendix C I establish that the solution to the individual bidding problem exists. The individual bidding problem can be reduced to the partial differential equation with boundary condition. In this way, the equilibrium can be reduced to the system of partial differential equations. As the proof in Appendix C still holds for vector-valued equations, it leads to the existence of the solution to such system of equations. Thus, it proves the existence of the equilibrium.

Uniqueness. I will now prove that the equilibrium is unique for the entire problem using the fact that the number of players in the game is driven by a Poisson shock, thus a situation when the number of bidders is extremely large is not possible with finite probability. I assume that the set of potential players in the game is countable. Problem (3.7) can be written for each player in the set of bidders $\mathbb{I}$.

Let us take finite subsamples of players from $\mathbb{I}$. For each fixed N - the size of a subsample drawn from the population optimization problems of players form a multidimensional boundary problem which has a unique solution for the same reason as a one-dimensional problem. It is known that the set of all finite subsets of a countable set is a countable set[3]. Therefore one can index each subsample of size N by some natural number k and attribute to it an element of a positive sequence $\{x_n\}_{n=1}^{\infty}$

[3]See [HJ99]

such that $\sum_{n=1}^{\infty} x_n = 1$. Further note that as players are coming at a Poisson rate one can find some positive constants c_1 and c_2 that the probability of encountering the subsample sample of size N is at least $c_1 e^{-c_2 N}$. Then the probability of observing a subsample of size N indexed k is at least $x_k\ c_1 e^{-c_2 N}$. Total probability of observing a subsample of size at most N is at least

$$\sum_{n=0}^{N} \sum_{k=1}^{\infty} x_k\ c_1 e^{-c_2 n} = \sum_{n=0}^{N} c_1 e^{-c_2 n} = 1 - e^{-(N+1)c_2}.$$

For subsamples of size N from $\mathbb{I}$ the above results establish the following properties:

(i) For each of the subsamples the equilibrium exists and is unique.

(ii) The total probability of the observing a subsample of a size at most N is approaching 1 as N goes to infinity.

When N goes to infinity, these two conclusions imply that with probability 1 the equilibrium exists and is unique.

This construction not only provides a method for calculating the equilibrium, but also provides results about the existence and uniqueness of the equilibrium.

This proof shows also that it is quite complicated to prove uniqueness in case when parameter set is infinite-dimensional (that is the relevant functions are defined nonparametrically). In fact to prove uniqueness in that case one would have to add some mechanism that exponentially increases explanatory power of the nonparametric estimate as the size of the grid increases.

3.3 Identification and Estimation

3.3.1 Identification

I consider an application of the model discussed in the previous section to a data-rich environment when multiple observations of a game with the same structure are available. Technically, this means that we observe multiple copies of the process x_t and a set of exogenous variables for each observation. Identification of the model for these data is determined by several factors. First, identification should depend on the number of types of players according to their utilities (this will determine how much utility functions can vary for different indices i, indicating particular players). Second, identification depends on the "tightness" of generalized market clearing conditions (3.4). For this reason, particular identification results will depend on the structure of a specific model under consideration. In this subsection we will illustrate identification of the model for a particular example of an auction game with two player types.

Example: Auction Game with Two Types of Bidders

The model above can be specialized to a dynamic auction game. For this game, the state variable x_t is a current price in the auction, p_t. In general, the action of the bidder is her bid b_t, but we are going to consider action $a_t = b_t - p_t$. Bidders are considered risk-neutral with utilities

$$u_v\left(a_T, p_t\right) = \left(v - p_t - a_t\right) \mathbf{1}\left\{a_t > 0\right\},$$

where v is the bidder's valuation. Suppose that variance of beliefs about θ takes either of two values $\sigma \in \{\sigma_1, \sigma_2\}$, where $\sigma_1 = 0$ and $\sigma_2 = +\infty$. In this case the beliefs of the first type of bidders coincide with the true value of visibility. The second type of bidders do not update their beliefs, as they have infinite variance at any point. In case of the second price auction the generalized market clearing condition for the Poisson-driven auction game can be written as

$$h\left(a_{it}, x_t, t\right) = \max_{k,\, k \neq \operatorname{argmax} a_{jt}} a_{kt}.$$

Properties of bidding behavior: Suppose that the distribution of valuations $F_v(\cdot)$ has a density $f_v(\cdot)$. The Bellman equation for the first type of bidders has a unique solution $b = p + a_t\left(t, p, v\right)$, where p is the current price in the auction and $a_t(\cdot)$ is the bid increment function. I index this bidding function by a subscript 1 to indicate that this is the bidding function of the first type of bidders. The solution of the Bellman equation for the first bidder type allows us to describe the properties of the solution in more detail. Specifically, if the instantaneous demand is decreasing in price and the price jump size function is appropriately bounded, then the bid will be monotone increasing in price. In this case one can invert the bidding function to express the price in terms of the valuation and the bid of the first type of bidders: $p = b^{-1}\left(b_1, v\right)$. Moreover, as the value function is increasing in the bidder's valuation, then the bidding function is also increasing in her valuation. If the price exceeds the bidder's valuation, then she stops bidding. For each t, the bid of the first type of bidder is an increasing function of price and bidder's valuation. As a result it is possible to express the distribution of bid of the first bidder type in terms of the distribution of price: $F_{b_1}(x) = \int_V F_{pt}\left(b^{-1}(x, v)\right) f_v(v)\, dv$. The bidders do not submit bids higher than their valuations, so that the probability mass for bids submitted by the bidders whose valuations are lower than the price is zero. As a result the distribution of bids of the first type of bidders can be written as:

$$F_{b_1}(x) = \int\limits_{x}^{+\infty} F_{p_t}\left(b^{-1}(x,v)\right) f_v(v)\, dv.$$

For the second type of bidder, the solution is also unique. The Bellman equation suggests that the value function for such bidders does not change over time. As a result, the optimal bidding for the second type of bidders implies that the bids are: $b = v$. The distribution of bids in this case will coincide with the distribution of valuations. In the following discussion, I use a subscript 2 to indicate that this bidding function corresponds to the second type of bidders. The distribution of bids of the second type of bidders is just the distribution of valuations: $F_{b_2}(x) = F_v(x)$.

Identifying parameter θ: Consider a collection of auctions for similar items which have started simultaneously. One may expect that the number of bidders will be different across these auctions. In my model there are two sources of variation in the number of bidders in such simultaneous auctions. The first source of variation is the variation in the price of the auction (including the starting price and the price movement over time). The second source of such variation is θ of the auction game. One can identify θ primely from differences across auctions. Consider the distribution of the number of active bidders across auctions given the price $F_N\left(\cdot\,|\,p_t,t\right)$. This distribution is determined by the integral variation of the price up to time t and the variation in θ across auctions. As a result, if one can remove the variation in the number of bidders due to the variation in the price it becomes possible to recover the distribution of θ. Specifically, in the beginning of the auction differences in parameters θ explain differences in the time until the first bid for auctions with the same starting prices. Proceeding with this argument, if one finds different auctions with "close enough" price trajectories, then the differences in the number of active bidders will be explained by the differences in θ. The proximity of price trajectories in different auctions is provided by the closeness of instantaneous demand functions in these auctions.

For each auction one can estimate the instantaneous demand as a frequency of price jumps $\widehat{\lambda}(t, p_t, \theta)$. To do this for a collection of auctions compute the number of price jumps per unit of time for different instances of the auction (by choosing some time grid and computing the number of jumps between the points on the grid or using a smoothing kernel) and "regress" these frequencies on the price and time variables. Informally, the residual of this "regression" in a specific auction will be associated with θ.

To be able to perform this step, I assume that θ are independent across auctions. Moreover, the instantaneous demand is monotone increasing in visibility and the support of θ is the segment $[0, 1]$[4]. In this case proceeding with the "regression" analysis of instantaneous demand, assign zero visibility to the auction with the smallest residual. Similarly we will assign value 1 to parameter θ in the auction with the highest residual in the instantaneous demand "regression". I assign θ to other auctions by scaling their residuals such that the scaled values are between 0 and 1. As a result, one can assign the values of θ to all auctions. This identification procedure relies on the normalization assumption for the domain of θ and the scale assumption for the instantaneous demand function. A different set of scaling assumptions will lead to different values for θ, but should preserve the rank of auctions according to their θ.

Identifying the distribution of valuations: For a fixed distribution of θ the distribution of valuation can be identified from the cross-section of bidders. Parame-

[4] In the theoretical model I assumed that θ takes values in [0, 1]. As a result the upper and the lower boundary of this set can be estimated by taking the maximum and the minimum of the estimated θ. In this example the instantaneous demand function takes the form $\lambda(t, p, \theta) = f(t, p) + a\theta$. Then taking into account that we did not impose the normalizing assumptions on the instantaneous demand function, one can use the normalization for θ to extract the parameter a. Under standard regularity conditions, such an estimator will be consistent.

ter θ affects the number of bidders through the instantaneous demand function. This means that the variation in the number of bidders induced by visibility will not depend on a particular normalization for the support of θ. In this case the analysis of a cross-section of auctions at each instant will be similar to the analysis of a collection of static auctions if the bidding functions are monotone.

If the distribution of bidder types is known, then the price in the auction will be determined by the mixture of bids submitted by bidders of two types. If bidding functions of both types of bidders are monotone with respect to their valuations[5], then the distribution of prices across auctions will be a monotone transformation of the distribution of valuations. In my example with two types of bidders, the distribution of prices in auctions at each instant is described by the distribution of the second highest bids. This is a second order statistic from the mixture of distributions $F_{b_1}(\cdot)$ and $F_{b_2}(\cdot)$. I have previously expressed these distributions in terms of the distribution of prices and the distribution of valuations. Given the proportion of the first type of bidders, one can invert the observed distribution of prices across auctions and obtain the distribution of valuations.

Identifying the bidder types: The data across time identify bidder types. The bids of the first type of bidders will have the conditional variance $var\,(p_t) + var\,(\eta\,(t, p_t, v))$. The second type of bidders will have bids with the conditional variance $var\,(v)$ (because they bid their valuations). As the bid increment function is decreasing in price and is non-increasing in valuation, bidders of the first type are identified as the bidders with lower variance of bids. In fact, if the distribution of val-

[5]The bidding function of the first type bidder will be monotone, for instance, when the instantaneous demand is bounded by some sufficiently small number when the price in the auction changes.

uations is known then one can exactly compute the variance of the bids of the second type of bidders. If some bidder has a variance of bid over time smaller than $var\,(v)$ then she will be automatically attributed to the first type. As the distribution of valuations is estimated with an error, then bidders can be sorted into types with the aid of a criterion which is equivalent to testing the null hypothesis: $var\,(b_t) = var\,(v)$. If the null hypothesis is rejected, then the bidder will be attributed to the first type.

3.3.2 Estimation setup

By the nature of the data-rich environment multiple sample paths of the same stochastic process x_t (given covariates) can be observed. In terms of the Poisson-driven model described above, this means that multiple sample paths of a Poisson process with frequency $\lambda\,(t, x, \theta)$ and jump size $h\,(t, x, a)$ are available. For estimation I assume that distribution of prices at each instant $t \in [0,\, T]$ given parameter θ is the same across auctions.

Assumption 3.1. *For any $t \in [0,\, T]$ in any given auction $x_t\,|\,\theta \sim F\,(t, \theta)$ and does not depend on a particular realization of the price process. Moreover, $F\,(t, \theta)$ is a proper distribution for each $t \in [0,\, T]$ and $\theta \in \Theta$.*

This means that I treat each state variable path as one single observation. In addition, I assume that these observations are independent and have the same finite-dimensional distributions. This can be considered as an extension of the notion of independent identically distributed observations to realizations of a stochastic process.

In the rest of the subsection I outline the estimation approach which is based on computing the optimal response of players to the state variable process. The intuition for the estimation approach used in this paper is outlined for a particular case of second-price auction in Chapter 1. Here I give a more formal treatment of the convergence results, by proving that the estimating functionals that we are using belong to the appropriate P-Donsker classes. My approach to estimation is close in flavor to the approach in [Gal04]. However, there is a major difference which makes my approach unique. In fact, due to assumption 3.1, in my framework one can observe multiple realizations of the same stochastic process. This means that unlike the data usually analyzed with indirect inference methods (such as EMM in [GT02]) which is a single realization of a stochastic process, it is possible to observe finite-dimensional distributions of stochastic process without imposing any stationarity restrictions.

I characterize the model of the game by a structural parameter $\gamma_0 \in \Gamma$, where Γ is some complete metric space. I assume that the model is identified by this structural parameter. This implies that given the true parameter value, it is possible to reproduce the distribution of state variables and beliefs at each instant of time by simulating the path of the game forward. Suppose now that k labels a particular game (i.e. a specific path of the state variable $x^{(k)}$), and index i will refer to a jump point of the state variable path, that is the state variable after the jump $x_i^{(k)}$ and the time of jump t_i. I normalize the duration of games to 1 so $t_i \in [0, 1]$. The total number jumps in the state variable in game k will be denoted I_k. Finally, let $N_t^{(k)}$ be the (unobservable) number of players in game k at the instant t.

The estimation procedure will follow 4 steps.

Step 1 Non-parametrically estimate the joint distribution of the state variable when it jumps and the time of its jumps. A simple and convenient method to do this when a significant amount of data is available is to estimate the joint density at $(x,\, t)$ by a kernel method:

$$\widehat{f}\,(x,t) = \frac{1}{n\,h_x\,h_t}\sum_{i=1}^{n}\sum_{i=1}^{I_k}\kappa\left(\frac{x_i^{(k)}-x}{h_x}\right)\kappa\left(\frac{t_i^{(k)}-t}{h_t}\right),$$

with the kernel function $\kappa\,(\cdot)$

Step 2 Fix structural parameter γ

Step 3 Rank auctions by simulated numbers of price jumps and estimate θ in each auction as a ratio of the rank of the auction and the total number of auctions. This provides an unbiased (but inconsistent) estimate of θ. Under certain conditions, population moments will be consistently estimated given these unbiased estimates.

Step 4 Simulate the number of players in each auction $N_t^{(k)}$ for any $t \in [0,\, T]$. For each player determine her utility function, and draw the initial mean and variance of the player's beliefs.

Step 5 For each player j compute the optimal strategy profile, which is the optimal action $a_{ij}^{(k)}\left(z_i^{(k)}\right)$ at the instant of jump in the state variable $t_i^{(k)}$. Then compute the "response" of the model to the data by solving the system (3.4) for unknown moments of the state variable and compute the implied movement in the state variable.

Step 6 Estimate the joint density of the model response to the observed state variable and the time of jumps, for instance, using kernel estimator:

$$\widehat{f}_\gamma\,(x,t) = \frac{1}{n\,h_x\,h_t}\sum_{i=1}^{n}\sum_{i=1}^{I_k}\kappa\left(\frac{\widehat{x}_i^{(k)}\,(\gamma)-x}{h_x}\right)\kappa\left(\frac{t_i^{(k)}-t}{h_t}\right).$$

Step 7 Compute the Kullback-Leibler information Criterion as

$$\widehat{KLIC}(\gamma) = \frac{1}{n} \sum_{i=1}^{n} \sum_{i=1}^{I_k} \log \left[\frac{\widehat{f}\left(x_i^{(k)}, t_i^{(k)}\right)}{\widehat{f}_\gamma\left(x_i^{(k)}, t_i^{(k)}\right)} \right], \tag{3.8}$$

and minimize it over γ by repeating steps 4 - 7.

The rationale for such estimation is that in equilibrium the optimal response to the state variable should replicate the changes in the state variable. Given a significant flexibility of the estimation method, it does not achieve the $\sqrt{n}$-convergence. However, the model is highly non-linear, and even to achieve a non-parametric convergence rate, we need to impose a list of restrictions on the data.

Assumptions

A 1 Realizations $\left\{x_\tau^{(k)},\, t_\tau^{(k)},\, N_\tau^{(k)}\right\}_{\tau\in[0,\,T]}$ are independent for each $\tau \in [0, T]$ and $k = 1, \ldots, n$, have joint and marginal distributions with finite second moments[6].

A 2 The kernel is a Borel-measurable function such that

- $\int \kappa(\psi) d\psi = 1$,
- $\int |\kappa(\psi)| d\psi < \infty$,
- $\kappa(\cdot)$ is twice continuously differentiable a.s. and approaches zero with its first and second derivatives as $|\psi| \to \infty$.
- $\sup_{\psi\in\mathbb{R}} |\kappa(\psi)| < \infty$

A 3 The marginal distribution function of the process $\left\{p_\tau^{(k)}, t_\tau^{(k)}\right\}_{\tau\in[0,\,T]}$ has a continuous Hessian which is bounded in some neighborhood of each point of the

[6]This assumption is based on the idea that in a large dynamic market, even though players can participate in multiple simultaneous games, they induce a very small correlation across the market.

domain.

A 4 $h_t,\ h_p \to 0$ as $n \to \infty$.

A 5 $n\, h_t\, h_p \to \infty$ and $n/\max\{h_t, h_p\}^2 \to 0$ as $n \to \infty$.

These assumptions ensure the appropriate convergence properties of the non-parametric density estimates. Moreover, these properties are helpful for the proof of asymptotic normality of the structural estimates. The following theorem establishes the convergence properties of the estimated densities and the properties of the structural estimates.

Theorem 3.2. *Suppose that Assumptions 1-5 hold. Then the density of the distribution of jumps in the state variable can be consistently estimated at step 1.*

In addition, assume that the frequency of jumps decays faster than the density of the state variable when the state variable takes large values

$$E\left\{\int_0^T \frac{\lambda(\tau, x_\tau)}{f^2(x_\tau, \tau, \gamma_0)}\, d\tau\right\} < \infty,$$

then the structural estimates obtained minimizing the objective defined at step 7 are asymptotically normal:

$$\sqrt{n h_t h_x}\,(\widehat{\gamma} - \gamma_0) \xrightarrow{d} N\left(0, Q^{-1}\,\Omega\, Q^{-1}\right),$$

where

$$Q = E\left\{\int_0^T \frac{\partial f(x_\tau,\tau,\gamma_0)/\partial\gamma}{f(x_\tau,\tau,\gamma_0)}\, dJ(\tau, x_\tau)\right\} \quad \textit{and} \quad \Omega = 2\left(\int_0^\infty \kappa^2(\psi)\, d\psi\right)^2 E\left\{\int_0^T \frac{dJ(\tau,x_\tau)}{f(x_\tau,\tau,\gamma_0)}\right\}.$$

The proof of this theorem is given in Appendix C.

Although Theorem 3.2 looks very similar to the standard asymptotic results in the non-linear settings, it is very different in the nature from the results derived for i.i.d. observations. In my model I employ the fact that different paths for the state variable are generated by the same stochastic process. Surprisingly, summation over observations as in (3.8) makes the problem equivalent to standard inference for "point" observations making (3.8) a function entire observed price paths. Individual components of (3.8) are independent and identically distributed across different paths. Theorem 3.2 establishes that the obtained estimates are asymptotically normal as the structural estimates with standard i.i.d. observations.

3.4 Welfare Comparisons of Dynamic Auction Mechanisms

3.4.1 Baseline model

In this section I apply the techniques that were discussed above for estimation of a general continuous-time dynamic model to a particular case of eBay auction. This model has been described in the Eeample given in Subsection 3.3.1[7]. Here I use the model to describe the welfare consequences of imposing a different set of auction rules (in particular, the rule used by Amazon on-line auction service).

3.4.2 Motivation of welfare analysis

Given the continuous-time structure of our model, it is optimal for bidders to deter entry by bidding early and to prevent learning by delaying their bids (until the last moment of the auction). These features suggest that the existing second-price auction structure with a fixed ending time may not be fully efficient. For example, in Chapter 1 it is demonstrated that an early bidder can raise the price and dissuade

[7]The details for estimation of this model can be found in Chapter 1, where the model is applied to estimation of eBay auctions for pop-music CDs.

potential bidders with high valuations from entering the auction. For this reason, it might be possible to adjust the eBay auction mechanism to increase the social welfare (the sum of ex ante expected surpluses of the buyers and the ex ante profit of the seller). This is the focus of the current section. The term ex ante describes the fact that we analyze the welfare of an average bidder (randomly drawn from the distribution of valuations). The analysis of ex ante expected welfare is based on the assumption that there is no correlation between the visibility of the auction and the valuations of the participating bidders.

My model allows me to approach the problem of optimal auction mechanism design from an applied perspective. Given the estimates of the primitives obtained from the actual eBay data, one can use my model to simulate counterfactual outcomes and evaluate the expected welfare of the participants in the auction under alternative auction designs. One such mechanism, in which the ending time of the auction depends on the arrival of new bids, was actually implemented by Amazon.com. The mechanism works as follows. After a certain time has elapsed from the beginning of the auction, the following mechanism is triggered: if there are no new bids within a fixed period of time the auction stops, otherwise the auction is extended for a fixed time period upon arrival of a new bid. This process continues until there are no new bids during the current extension period. Both the extension time and the instant when the extension mechanism starts are known to all bidders.

This mechanism could lead to welfare gains, since it may reduce the incentives for late bidding, which bidders on eBay can use to prevent learning by other bidders. However, as I demonstate in this section, this is not actually the case: although there is a positive welfare gain from switching from the eBay to the Amazon auction mechanism, it is generally quite small. The welfare gain becomes significant only when the extension time on Amazon is very large. A long extension time for the Amazon auction is not realistic because the bidders also have a cost of time which reduces the gain from winning the auction. In this case, when the duration of the auction is very long the bidders might choose to buy the auctioned object elsewhere.

In the following subsection, I adapt my analytical framework to nest the Amazon structure. Because the Amazon auction can be extended indefinitely (if new bids keep arriving) the potential bidder problem must be solved out to infinity. To compute the welfare changes I use the structural estimates for eBay auctions for pop-music CDs on eBay obtained in Chapter 1. Using these estimates I compute the strategies and the welfare of the bidders on an Amazon auction by solving the bidder's problem for the new objective function.

3.4.3 Continuous-time Amazon auction model

In this section I use the continuous-time techniques which can be applied to the analysis of eBay auctions to analyze the mechanism of Amazon. First of all, I describe the continuous-time auction mechanism which I call "the Amazon mechanism". The current price in the auction is determined as the second highest bid of all the bids submitted up to the current moment. The auction starts at time 0 and can potentially continue to infinity. However, at time T the following mechanism is triggered. If there are no more bids submitted in the interval $[T,\, T+\tau]$ then the auction ends at time $T+\tau$ and the winner is the highest bidder at time T. If there is a bid which arrives at the moment $T < t' < T+\tau$ then the auction extends for time τ from the moment of arrival of the last bid. Again, if there are no more bids during the extension time τ (i.e. no bids in the interval $[t',\, t'+\tau]$) then the auction ends at the end of the extension period at the moment $t'+\tau$ and the winner is the highest bidder at the moment t'. If a new bid arrives during the extension period, the auction extends again for the same amount of time τ. This process continues until there are no more bids arriving during the extension period.

Suppose that a particular bidder is the winner of the Amazon auction and the auction ends at some moment $t+\tau$ where $t>T$. By the aforementioned rules, this means that this bidder was the highest bidder at the moment t and no new bids arrived during the period from t to $t+\tau$. Note that as the auction has to continue up to the moment $T+\tau$ in any case, then the fact that the auction ends at time $T+\tau$ and the considered bidder is the winner, means that the new bids stopped arriving at some moment $t<T$ and at that moment the considered bidder was the highest bidder.

In terms of the counting process of price jumps $J(t,p_t,\theta)$ the fact that the auction ends at time $t+\tau$ $(t>T)$ and, thus, there are no bids arriving from time t to time $t+\tau$ means that $\int_t^{t+\tau} dJ(r,p_r,\theta)=0$, which means that the number of new bids is equal to zero. If the auction ends at time $T+\tau$, the new bids stop arriving at some moment $t<T$ and thus the number of new bids $\int_t^{T+\tau} dJ(r,p_r,\theta)=0$. The moment ξ after which there are no bids up to moment $T+\tau$ (if $\xi<T$) or no bids up to the moment $\xi+\tau$ (if $\xi>T$) can be expressed as a minimal time after which the bids stop arriving and the auction ends. Note that if we know ξ then we know the moment when the auction ends. Formally this instant can be expressed as[8]:

$$\xi=\min\left\{\inf_{t\leq T}\left[\int_t^{T+\tau} dJ(r,p_r,\theta)=0\right] \text{ and } \inf_{t>T}\left[\int_t^{t+\tau} dJ(r,p_r,\theta)=0\right]\right\}.$$

If some bidder is the highest bidder at the moment ξ then she is the winner of the

[8]I assume that if $arg\ \inf(\cdot)=\emptyset$ then $\inf(\cdot)=+\infty$. This assumption is standard and it allows me to avoid considering different cases and write the time ξ in a parsimonious form.

auction (because there are no new bids after the moment ξ). The probability that a certain moment is the "winning time" $t = \xi$ is equal to the probability that there are no bids in the interval $[t,\, T + \tau]$ if $t < T$ and it is equal to the probability that there are no bids in the interval $[t,\, t + \tau]$ if $t > T$. This probability is closely related with the structure of the entry into the auction.

Denote $\Lambda_t(\Delta t) = \int\limits_t^{t+\Delta t} \lambda\left(r,\, p_r,\, \theta\right)\, dr$, which is a cumulative instantaneous demand from the moment t to the moment $t + \Delta t$. This cumulative instantaneous demand reflects the expected number of bids that can arrive in the auction during the interval Δt conditional on the price at the instant t. The indicator that a certain instant is earlier than the instant of the last bid ξ has an expectation

$$E_t\left[\mathbf{1}\left\{\xi > t\right\}\right] = \gamma(t) = \begin{cases} 1 - e^{-\Lambda_T(T+\tau-t)} & \text{if } t < T, \\ 1 - e^{-\Lambda_T(\tau)} & \text{if } t \geq T. \end{cases}$$

This expression is due to the fact that the bid arrival process is Poisson. In this way the model becomes similar to the conventional survival model, where the risk of failure is measured by the probability that there are no new bids after the moment ξ. The distribution of the moment ξ conditional on the price at this moment is exponential with cumulative hazard rate $\Lambda_t(\cdot)$. The conditional expectation written above describes the probability of "survival" at the moment t. Then, in order co compute the utility of the bidder from the auction, note that the bidder wins the object and pays the price p_t if $t = \xi$, and the bidder's bid increment is $\eta > 0$. Otherwise, the auction will continue. Then if t_i are the moments of price jumps (index i denotes a particular price jump) and h_i are the sizes of price jumps at the moments t_i then the utility of the bidder from winning the auction can be expressed as:

$$U = u\left(v - p_0\right) + \left[u\left(v - p_0 - h_0\right) - u\left(v - p_0\right)\right] \mathbf{1}\left\{t_1 < \xi\right\} \\ + \ldots + \left[u\left(v - p_k - h_k\right) - u\left(v - p_k\right)\right] \mathbf{1}\left\{t_{k+1} < \xi\right\} + \ldots \tag{3.9}$$

Note that if t_n is equal to ξ then all the indicators $\mathbf{1}\left\{t_k < \xi\right\}$ at the the instants of price jumps earlier than the instant t_n will be equal to 1 and all the indicators for the moments of price jumps later than t_n will be equal to 0. As a result the total utility of the bidder in the case of positive bid increment at the moment t_n is

$$U = u\left(v - p_0\right) + \left[u\left(v - p_0 - h_0\right) - u\left(v - p_0\right)\right] + \ldots + \left[u\left(v - p_k - h_k\right) - u\left(v - p_k\right)\right] \\ + \ldots + \left[u\left(v - p_n - h_n\right) - u\left(v - p_n\right)\right] = u\left(v - p_n - h_n\right).$$

This is exactly the ex post utility from winning the auction. Note that the moments of price jumps $t_k > t_n$ reflect the potential price jumps. They arrive too late so that the extension time τ is elapsed and the auction is closed. The summation in the suggested formula for the utility of the bidder goes over the price jumps, as a result, the ex ante expected utility of the bidder can be written in a compact form as:

$$V_0 = E_0\left\{U\right\} \\ = u\left(v, p_0\right) + E_0\left\{\int_0^{+\infty} \left[u\left(v,\, p_t + h(t, p_t, a_t)\right) - u\left(v,\, p_t\right)\right] \mathbf{1}\left\{t < \xi\right\}\, dJ_\epsilon(t,\, p_t,\, \theta)\right\}.$$

Assuming that the necessary regularity conditions hold, one can compute the expectation of the expression in the integral given the information at time t. In this case the expectation of the indicator function $\mathbf{1}\left\{t < \xi\right\}$ given the information at time t is the probability that the auction will not end at time t. The probability of the auction continuing is the survival probability with the cumulative hazard rate $\Lambda_t\left(\Delta\, t\right)$, and the expression for this probability was obtained above. Using this expression for the conditional expectation of the indicator function and the conditional expectation for the increment of the Poisson process we obtain:

$$V_0 = u\left(v, p_0\right) + E_0 \left\{ \int\limits_0^{+\infty} \gamma(t) \left[u\left(v,\, p_t + h(t, p_t, a_t)\right) - u\left(v,\, p_t\right) \right] \lambda_\epsilon(t,\, p_t,\, \theta)\, dt \right\}.$$

If bidders are risk neutral then price jumps lead to a decrease in their utilities. The formula suggests that the utility of the bidder will be decreasing from the level $u\left(v, p_0\right)$, corresponding to the bidder's surplus if she could buy the object for the initial price, at the rate $\lambda_\epsilon(\cdot)$ corresponding to instantaneous demand of the bidder, weighted by the survival probabilities. This means that the utility of the bidder decreases faster over time in the auctions with the higher instantaneous demand. However, the decrease in utility over time is compensated by the increase in the probability of the auction stopping. This probability is determined by the cumulative instantaneous demand in the auction.

In these circumstances the instantaneous demand determines both the utility of the bidder and the instant of time when the auction ends. If the bidder learns about the visibility of the auction, she also simultaneously learns about the expected time when the auction ends (which can be retrieved from the distribution of the moment ξ). As a result, the optimal learning equations for the bidder in the Amazon auction will be the same as for the bidder in the eBay auction. The Amazon bidder solves:

$$\begin{aligned} &\max_{a_t(t,p,\theta)} E_0 \left\{ U \right\} = E_0 \left\{ \int_0^{+\infty} \gamma(t) \left[u\left(v,\, p_t + h(t, p_t, a_t)\right) - u\left(v,\, p_t\right) \right] \lambda_\epsilon(t,\, p_t,\, \theta)\, dt \right\} \\ &dp_t = h(t,\, p_t,\, a_t)\, dJ_\epsilon(t,\, p_t,\, \theta), \end{aligned}$$

along with the optimal learning equations. This optimization problem can be used to form a stochastic Bellman equation. Defining the value function as the expected utility at time t, the Hamilton-Jacobi-Bellman equation can be written as:

$$\begin{aligned}
&\tfrac{\partial V_a(t,p,\mu,\sigma)}{\partial t} + \sup_{a_t \in \mathcal{A}} \Big[-\tfrac{\sigma^2}{\lambda} \left(\tfrac{\partial \lambda}{\partial \theta}\right)^2 \tfrac{\partial V_a(t,p,\mu,\sigma)}{\partial \sigma} - V_a(t,\, p,\, \theta,\, \sigma)\, \lambda_\epsilon\left(x,\, \theta,\, t\right) + \\
&+ V_a\left(t,\, p + h\left(t, p, a_t\right),\, \mu + \tfrac{\partial \lambda}{\partial \theta}\tfrac{\sigma}{\lambda},\, \sigma\right) \lambda_\epsilon\left(p + h\left(t, p, a_t\right),\, \mu + \tfrac{\partial \lambda}{\partial \theta}\tfrac{\sigma}{\lambda},\, t\right) + \\
&+ \gamma(t)\left[u\left(v,\, p_t + h(t, p_t, a_t)\right) - u\left(v,\, p_t\right)\right] \lambda_\epsilon(t,\, p_t,\, \mu) \Big] = 0
\end{aligned} \tag{3.10}$$

$$\lim_{t \to +\infty} V_a(t,\, p,\, \mu,\, \sigma) = u\left(v,\, p\right)$$

There are two major differences between the first order conditions for the bidders on Amazon and on eBay. The first difference is that, for the eBay bidder, the utility enters into the first order condition only once - as a boundary value of the value function. On Amazon, the utility determines the value function at each instant. This is because on eBay the bidder knows exactly when the auction ends and thus cares only about the utility at that moment while on Amazon the ending time is uncertain and the bidder should take into account that with a certain probability each moment can be the ending time of the auction[9]. The second difference is that unlike the bidder's problem on eBay, the bidder's problem on Amazon only has an asymptotic boundary condition and the reason for this is the Amazon auction ends with certainty only at infinity.

The expected utility of the seller can be obtained from the same expressions as the expected utility of the bidder. The only difference is that the instantaneous (static) utility of the seller is increasing in the sale price of the auctioned object. The overall

[9]I define the "ending" time ξ as the instant after which new bids do not arrive. Therefore, it can occur even before the fixed duration length T.

utility of the seller is the integral of the the instantaneous utilities discounted by the auction extension probability $\gamma(\cdot)$. If $u^*(v^*, p)$ determines the instantaneous utility of the seller from selling the object on Amazon, then the seller's problem is:

$$\max_{a_t} E_0 \left\{ \int_0^{+\infty} \gamma(t) \left[u^* \left(v^*, \, p_t + h^*(t, p_t)\right) - u^* \left(v^*, \, p_t\right)\right] \lambda_\epsilon(t, \, p_t, \, \theta) \, dt \right\}$$
$$dp_t = h^*(t, \, p_t) \, dJ_\epsilon(t, \, p_t, \, \theta).$$

The solution to the seller's problem is expressed by the Hamilton-Jacobi-Bellman equation, similar to the solution of the bidder's problem.

The main difference between the optimal bidding and selling problems on eBay and Amazon is that the instantaneous utility of the Amazon auction participants explicitly influences the decision at each point of time. Moreover, the influence of the boundary condition (stating that at the end of the auction the utility of the bidder is equal to the surplus from winning the auction) is significantly smaller on Amazon than on eBay. This is because it has to be satisfied asymptotically on Amazon.

3.4.4 Computing the welfare on eBay and Amazon

I define the *ex ante* social welfare in a continuous-time auction as the sum of expected surplus of buyers and the expected surplus of the seller. To compute the welfare of the buyers at the instant they enter it is necessary to compute the expected surplus of the bidders over all possible numbers of entrants. For a given number of entrants the expectation is taken over the possible valuations of the entrants.

Thus if $V(\cdot)$ is the value function of the bidder computed from the optimal bidding problem, then the *ex ante* expected surplus of the bidders is:

$$W = E \left\{ \int_0^{+\infty} \int_0^{T} V(t, \, p_t, \, \mu_0, \, \sigma_0, \, v) \, dJ_\epsilon \left(t, p_t, \theta_0\right) dF_v(v) \right\},$$

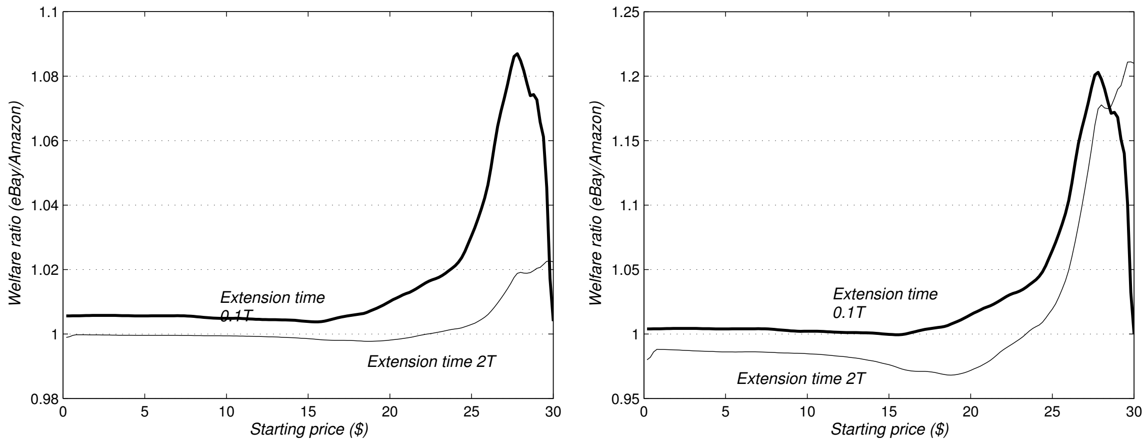
where T is set to infinity for the Amazon mechanism, and $F_v(\cdot)$ is the distribution of valuations. For comparison I computed the welfare ratios for the bidders with perfect information about the visibility of the auction and for the bidders who are learning about the visibility.

In computations I use the structural estimates obtained in Chapter 1 which characterize the structure of the entry into the auctions as well as parameters of distributions and beliefs. To compare the surplus of the bidders in the eBay and the Amazon auctions I computed the expected surplus of the bidders as a function of the starting price based on the results of the structural estimates of the bidding parameters for pop-music CDs on eBay. During the computations, the distribution of valuations is considered to be non-standard chi-square as in the estimation procedure, with the parameters estimated using MCMC. The integral over the distribution of valuations was computed using Monte Carlo integration techniques. Monte-Carlo integration was performed by sampling 1000 valuations from the distribution of valuations. For each valuation, the expected surplus from winning the auction was computed and then the result was averaged across the sampled valuations. If the size of the sample is large enough, then the result of such averaging should converge to the integral under consideration. Similarly, for the bidders learning about the visibility of the auction in addition to valuation, I drew values for the initial beliefs of the learning bidders and then averaged the results over draws.

For welfare comparisons, the welfare of the bidders in Amazon auctions was computed for the extension time τ equal to $0.1\,T$ and $2\,T$ where T is the duration of the auction (when the extension time mechanism is triggered). Figure 3.1 displays the ratio of the welfare of the bidders on eBay to the welfare of the bidders on Amazon.

As one can see from Figure 3.1, the welfare on Amazon tends to be smaller than

Figure 3.1: The ratio of the expected surplus of the bidder on eBay to the surplus on Amazon auction



(a) Learning bidders (b) Experienced bidders

the welfare on eBay if both the extension time τ and the starting price is small. However, if the Amazon auction extends for a long time after each subsequent price jump (which means that the extension time τ is large) the welfare of the bidders tends to be higher on Amazon than on eBay for a broad range of starting prices. The intuition for this result is that once the Amazon auction extends for a long time, the bidders gain control over the auction ending time by submitting predatory bids. This additional degree of freedom allows the bidders to improve their welfare in the Amazon auction relative to the eBay auction. The welfare of the bidders on Amazon is lower than on eBay if the extension time is small. The rationale for this result is that, in this case, the increment of the bidders' welfare from the added flexibility of the stopping time does not completely offset the decrease in welfare due to the possibility of increased entry towards the end of the auction.

Comparison of panels (a) and (b) shows that the differences in the welfare on eBay and on Amazon are smaller for the bidders who do not know the visibility of the auction perfectly. This suggests that the absence of information leads to the loss of bidder's control over the ending of the auction in the Amazon auction case and the welfare of the bidders on eBay and on Amazon become very similar.

3.5 Conclusion

In this paper, I design a framework for modelling and estimation of continuous-time interactions. I extend the continuous-time model in Chapter 1 and establish the existence and uniqueness of the dynamic equilibrium. I then outline the identification argument for the specific case of model, which we use to build an estimator. Further, I formally set up the estimator and prove its convergence.

The proposed strategy turns out to be a convenient tool for the analysis of consumer's welfare and seller's profit impact of policy changes. Using the structural estimates for eBay auctions obtained in Chapter 1 I compare the social welfare in the case of the continuous-time auction for eBay auction design and for Amazon auction design where the ending of the auction is conditional on the arrival of new bids. Computations show that depending on the parameter values, the Amazon auction can either enhance predatory bidding behavior or diminish it. The latter is likely to occur if after the arrival of a new bid the auction continues for a sufficiently short interval of time.

Appendix A

Tables for Empirical Results in Chapter 1

Table A.1: Summary statistics for auctions for Madonna's CDs

Variable	*N. obs*	*Mean*	*Std. deviation*	*Min*	*Max*
		Auctions			
Buy-it-now price	249	14.7676	46.7879	0.01	577
Starting bid	1132	5.0984	9.2041	0.01	175
Condition	1281	0.5097	0.5001	0	1
Picture	1281	0.9633	0.1880	0	1
Duration	1281	6.5011	1.3270	1	10
Store	1281	0.3575	0.4794	0	1
% seller's positive feedback	1281	98.6860	8.6711	10	100
Seller's feedback score	1281	9131.411	34726.98	0	325614
		Bidding			
time	1591	0.7775	0.3010	0.0015	1
bid	1591	13.8287	28.3903	0.01	699

Table A.2: Results of regression analysis.

Dependent variable	*price jump*	*bids per time*	*number of bids*	*number of bidders*
Regressor	(1)	(2)	(3)	(4)
Time	-16.141	-	-	-
	(-6.43)**			
Bid	1.152	-	-	-
	(38.36)**			
Starting price	-	-0.091	-0.004	-0.005
		(-2.17)**	**(-2.11)***	**-1.86**
Confessions CD	-	10.27	0.513	1.075
		(3.10)**	**(2.80)****	**(2.45)***
Confessions Tour	-	32.263	1.197	1.642
CD+DVD		**(3.22)****	**(3.37)****	**(3.07)****
Duration	-	-6.224	-0.003	-0.033
		(-4.15)**	**-0.09**	**-0.54**
Picture	-	2.644	0.154	0.158
		0.92	**0.87**	**0.41**
Shipping cost	-	0.625	0.008	0.096
		-1.1	**-0.31**	**-1.6**
% seller's positive	-	-0.035	-0.003	-0.002
feedback		**-0.34**	**-0.54**	**-0.19**
Seller's feedback	-	1.232	0.072	0.118
score		**(5.17)****	**(6.10)****	**(4.27)****
Constant	-0.752	50.528	0.866	0.878
	-0.37	**(3.52)****	**-1.35**	**-0.81**
Observations	1591	1280	1280	1280
F-statistic	737.6	28.78	7.64	4.75

t-statistics are given in the parenthesis. Coefficients for which standard errors are marked by ** are significant at 1% confidence level marked by * are significant at 5% confidence level.

Table A.3: Response to changes in beliefs and the instantaneous demand

Changes in average variance of beliefs $E[\sigma_\theta]$				
Factor	0.5	1.0	1.5	2.0
Fraction of early	2.1%	3.4%	7.5%	9.4%
bids	(0.8)	(1.1)	(3.2)	(7.0)
Average ratio of bid and	12.0%	12.2%	7.9%	10.0%
the final price	(6.0)	(5.2)	(3.3)	(5.5)
Changes in the coefficient α_λ				
Factor	0.5	1.0	1.5	2.0
Fraction of early	2.2%	3.4%	6.0%	10.2%
bids	(1.1)	(1.5)	(2.3)	(1.1)
Average ratio of bid and	10.0%	12.2%	13.8%	16.7%
the final price	(3.9)	(5.2)	(7.4)	(9.0)
Changes in the coefficient a_1				
Factor	0.5	1.0	1.5	2.0
Fraction of early	4.3%	3.4%	7.0%	10.0%
bids	(1.8)	(1.1)	(5.1)	(7.2)
Average ratio of bid and	8.8%	12.2%	16.0%	15.7%
the final price	(5.0)	(5.2)	(6.5)	(8.0)

Standard errors are in the parentheses

Table A.4: Characteristics of the auctions for "Greatest hits" CDs in the control group

variable	N obs.	mean	stdev	min	max
Buy it now price (US $)	80	14.97	17.60	1.89	63.72
Picture dummy	136	0.8823529	0.3233808	0	1
Duration (days)	136	7.632353	1.919911	3	10
Shipping cost ($)	136	7.269701	6.490984	0	37.2
Seller's feedback score	136	13414.46	22421.4	0	164546
Store dummy	136	0.6029412	0.4910972	0	1
Condition (new=1)	136	0.6911765	0.4637162	0	1
Number of bids	136	1.875	3.009707	0	10
Number of bidders	136	1.279412	1.95388	0	8
Average bid	53	3.984911	3.505953	0.01	25.04667

Table A.5: Characteristics of the bidders for "Greatest hits" CDs in the control group

variable	N obs.	mean	stdev	min	max
Feedback score	230	209.7217	557.9766	0	4611
% positive feedback	230	99.63739	1.183508	92.6	100
Bought CD last 3 months	230	0.4565217	0.4991924	0	1
Bought from eBay	230	0.9130435	0.2823859	0	1
Europe dummy	230	0.8521739	0.3557016	0	1
Australia dummy	230	0.0652174	0.2474476	0	1

Table A.6: Characteristics of the auctions for "Greatest hits" CDs in the treatment group

variable	N obs.	mean	stdev	min	max
Buy it now price (US $)	45	7.98	7.33	2.98	34.37
Picture dummy	156	0.8205128	0.3849957	0	1
Duration (days)	156	7.685897	2.23406	3	10
Shipping cost ($)	156	4.924754	2.894402	1.59	18.1
Seller's feedback score	156	34448.39	65604.64	0	176262
Store dummy	156	0.4615385	0.5001241	0	1
Condition (new=1)	156	0.9102564	0.2867346	0	1
Number of bids	156	0.6858974	1.751698	0	12
Number of bidders	156	0.5192308	1.127207	0	7
Average bid	46	3.393384	3.278704	0.98	14.99

Table A.7: Characteristics of the bidders for "The Greatest hits" CDs in the treatment group

variable	N obs.	mean	stdev	min	max
Feedback score	102	276.4412	716.481	0	4795
% positive feedback	102	99.64902	1.370751	87	100
Bought CD last 3 months	102	0.4313725	0.4977137	0	1
Bought from eBay	102	0.9117647	0.2850375	0	1
Europe dummy	102	0.6568627	0.4771014	0	1
Australia dummy	102	0.0980392	0.2988362	0	1

Table A.8: Effect of market expansion on early bidding: dependent variable - early bidding dummy

	Threshold		
Variable	10%	15%	20%
Treatment dummy	0.6379	0.8134	0.7091
	(0.2776)**	**(0.2980)*****	**(0.2835)****
Condition	-0.9041	-0.9609	-0.7330
	(0.2960)***	**(0.3053)*****	**(0.3040)****
Duration	-0.0914	-0.0895	-0.0840
	(0.0545)	**(0.0541)***	**(0.0556)**
Shipping cost	0.0024	0.0141	-0.0015
	(0.0221)	**(0.0195)**	**(0.0247)**
Seller's feedback score	-0.0053	-0.0071	-0.0098
	(0.0260)	**(0.0257)**	**(0.0259)*****
Store dummy	-0.5876	-0.5088	-0.3891
	(0.2526)**	**(0.2533)****	**(0.2486)**
Constant	-0.0011	-0.2262	-0.3783
	(0.4715)	**(0.4588)**	**(0.4677)*****
N. obs.	292	292	292
pseudo - $\mathbf{R}^2$	0.1530	0.1637	0.119

Table A.9: Multiple bidding as a function of experience

Variable	Model (1)	(2)	(3)
Treatment	.0968	.0306	.0298
	(.0704)	**(.0980)**	**(.0528)**
Bought CD last 3 months	-.0421	-	-
	(.0218)*		
Bought CD × Treatment	.1218	-	-
	(.0430)***		
Bought from eBay	-	-.0870	-
		(.0411)**	
Bought from eBay × Treatment	-	0.1351	-
		(.0642)*	
Feedback score	-	-	-.4969
			(.1577)**
Feedback score × Treatment	-	-	.4702
			(.2556)*
Europe dummy	-.2937	-.3042	-.3013
	(.0533)***	**(.0521)*****	**(.0528)*****
Australia dummy	.-2336	.-2497	-.2578
	(.0691)***	**(.0670)**	**(.0662)**
Constant	.6143	.6845	.6114
	(.0551)***	**(.0640)*****	**(.0537)*****
N. obs.	332	332	332
$\mathbf{R}^2$	.2361	.2359	.2580

Appendix B
Proofs and Tables in Chapter 2

B.1 Approximation properties of the simulation algorithm

Let us first analyze the precision of approximation of the variables D_T, W_T and Z_T if the structural parameters in the simulations correspond to the structural parameters of the data generating process. Let $x^*(t, D, \theta)$ be the optimal strategy of the manager given the time, market demand and consumers; tastes. Consider then the expectation of the difference between the simulated and the actual stochastic processes:

$$E\left\|D_T - D_T^{(s)}\right\|^2 = E\left\|\int_0^T \psi(t, D_t, x^* - \theta_t)\,dt + \int_0^T \zeta(t, D_t, x^* - \theta_t)\,dB_t^1 - \right.$$
$$\left. - \sum_{i=1}^{N_\tau}\left\{\psi(t_i, D_{t_i}^{(s)}, x^* - \theta_{t_i})\tau + \zeta(t_i, D_{t_i}^{(s)}, x^* - \theta_{t_i})\Delta\,\widehat{B}_{t_i}^1\right\}\right\|^2.$$

If the functions $\psi(\cdot)$, $\zeta(\cdot)$ and $\gamma(\cdot)$ are continuously differentiable and bounded from above by some constant C in their domains then there exist constants $L_i > 0$ such that:

$$E\left\|\int_0^T \psi(t, D_t, x^* - \theta_t)\,dt + \int_0^T \zeta(t, D_t, x^* - \theta_t)\,dB_t^1 - \right.$$
$$\left. - \sum_{i=1}^{N_\tau}\left\{\psi(t_i, D_{t_i}^{(s)}, x^* - \theta_{t_i})\tau + \zeta(t_i, D_{t_i}^{(s)}, x^* - \theta_{t_i})\Delta\,\widehat{B}_{t_i}^1\right\}\right\|^2 \leq$$
$$\leq L_1 \sum_{i=1}^{N_\tau} E\left\{\sup_{t\in[t_i,\, t_{i-1}]} \|D_t - D_{t_i}^s\|^2\right\} + L_2 \sum_{i=1}^{N_\tau} E\left\{\sup_{t\in[t_i,\, t_{i-1}]} \|\theta_t - \theta_{t_i}^s\|^2\right\} \leq$$
$$\leq 2\,(L_1 + L_2)\,C^2 N_\tau \tau^2.$$

For the uniform choice of the time grid $N_\tau = T/\tau$ therefore, $E\left\|D_T - D_T^{(s)}\right\|^2 \to 0$ as $\tau \to 0$. We can provide the similar conditions to assure the convergence of the simulated variables $W_T^{(s)}$ and $Z_T^{(s)}$.

I obtained that the simulated paths of the considered stochastic process will have the moments close enough to the moments of the actual stochastic process. We should now establish the same result for the distribution of the simulated response and the actual process. For this reason consider the distribution function of the actual random variable $\mathbf{P}\left\{D_T < \delta,\, W_T < \omega,\, P_t < \pi\right\}$ and the distribution of the simulated random variable $\mathbf{P}\left\{D_T^{(s)} < \delta,\, W_T^{(s)} < \omega,\, P_t^{(s)} < \pi\right\}$ and consider the difference between these distribution in the fixed point $(\delta,\, \omega,\, \pi)$. For the distance between two distributions we can write the following chain of expressions:

$$\begin{aligned}
&\left| \mathbf{P}\left\{D_T < \delta,\, W_T < \omega,\, P_t < \pi\right\} - \mathbf{P}\left\{D_T^{(s)} < \delta,\, W_T^{(s)} < \omega,\, P_t^{(s)} < \pi\right\}\right| = \\
&\Bigg| E\left[\mathbf{1}\left\{D_T < \delta\right\}\mathbf{1}\left\{W_T < \omega\right\}\mathbf{1}\left\{P_T < \pi\right\}\right] \\
&\qquad\qquad -E\left[\mathbf{1}\left\{D_T^{(s)} < \delta\right\}\mathbf{1}\left\{W_T^{(s)} < \omega\right\}\mathbf{1}\left\{P_T^{(s)} < \pi\right\}\right]\Bigg| \leq \\
&\leq \left| E\left[\varphi^{\gamma}\left(D_T,\, W_T,\, P_T\right) - \varphi^{\gamma}\left(D_T^{(s)},\, W_T^{(s)},\, P_T^{(s)}\right)\right]\right|,
\end{aligned}$$

where the function $\varphi^{\gamma}(\cdot)$ approaches the indicator function as $\gamma \to 0$ and is absolutely integrable and infinitely differentiable with respect to its arguments. Assuming that the support of $\varphi^{\gamma}(\cdot)$ corresponds to that of the distribution of the random vector $(D_T,\, W_T,\, P_T)$, then the expectation of the derivative $\varphi^{\gamma}(\cdot)$ has a finite limit. As a result we can evaluate the distance from above:

$$\begin{aligned}
&\left| \mathbf{P}\left\{D_T < \delta,\, W_T < \omega,\, P_t < \pi\right\} - \mathbf{P}\left\{D_T^{(s)} < \delta,\, W_T^{(s)} < \omega,\, P_t^{(s)} < \pi\right\}\right| \leq \\
&\leq C_1^{\gamma} E\left\|D_T - D_T^{s}\right\| + C_2^{\gamma} E\left\|W_T - W_T^{s}\right\| + C_3^{\gamma} E\left\|P_T - P_T^{s}\right\| \leq \Lambda^{\gamma}\tau\sqrt{N_{\tau}}.
\end{aligned}$$

The last inequality follows from the relationships established for each of the components of the considered stochastic process. The previous inequality follows from differentiability of $\phi^{\gamma}(\cdot)$ and the fact that $\int \left|\phi^{\gamma}\left(x,y,z\right)\right| dx\, dy\, dz < \infty$. Thus, we have established the result that the distribution function of the simulated stochastic process approaches the distribution function of the observed stochastic process. If the distribution functions are assumed to be smooth, then the density $\widehat{f}_{\theta}^{(s)}(D, W, P)$ is converging to the density of $\left(D_T^{(s)}, W_T^{(s)}, P_T^{(s)}\right)$.

It has been also shown that the rate of the mean square convergence is at least $\sqrt{\tau} \sim N_\tau^{-1/2-\psi}$, for $\psi > 0$. As a result, the simulation error will be negligible as compared with the density estimation error if the size of the simulation grid N_τ significantly exceeds the value $nh_D h_w h_p$ determining the rate of non-parametric convergence of the density estimate.

A7 There exists some $\psi > 0$ such that $N_\tau^{-\psi} n h_D h_w h_p \to 0$ as $n \to \infty$.

In fact we will need to provide the asymptotic distribution of the non-parametric estimate of the density of the simulated sample. We can provide the following chain of expressions:

$$
\begin{aligned}
&\sqrt{n\,h_D\,h_w\,h_p}\left[\widehat{f}_{\theta}^{(s)}(D,\,W,\,P) - f_\theta\,(D,\,W,\,P)\right] \\
&\sqrt{n\,h_D\,h_w\,h_p}\left[\widehat{f}_{\theta}^{(s)}(D,\,W,\,P) - f_\theta^{(s)}\,(D,\,W,\,P)\right] \\
&\qquad\qquad + \sqrt{n\,h_D\,h_w\,h_p}\left[f_{\theta}^{(s)}(D,\,W,\,P) - f_\theta\,(D,\,W,\,P)\right] \\
&= \sqrt{n\,h_D\,h_w\,h_p}\left[\widehat{f}_{\theta}^{(s)}(D,\,W,\,P) - f_\theta^{(s)}\,(D,\,W,\,P)\right] + \Lambda\,N_\tau^{-\psi}\sqrt{n\,h_D\,h_w\,h_p}.
\end{aligned}
$$

The last component of the presented expression vanishes due to the assumption **A7**. For the first component we can apply a standard result from the kernel density estimation. It suggests that under assumptions **A1** - **A6** the density estimate is asymptotically normal. As a result for the whole expression we can write:

$$\sqrt{n\,h_D\,h_w\,h_p}\left(\widehat{f}_\theta^{(s)}(D,\,W,\,P)-f_\theta(D,\,W,\,P)\right)$$
$$\xrightarrow{d} N\left(0,\,f_\theta(D,\,W,\,P)\left\{\int\limits_{-\infty}^{+\infty}\kappa^2(x)\,dx\right\}^3\right),$$

given the assumption imposed on the bandwidth parameters and kernel function and an additional assumption that $nh_Dh_wh_p/N_\tau \to 0$ as $n \to \infty$.

One of the specific problems of the used approach is that the empirical density of the model response for different parameter values will be smooth with respect to the structural parameter only in the limit. In general, if we rely on the likelihood inference there is a need in computing the derivative of the density with respect to the structural parameters. To simplify the expressions denote $x = (D,\,W,\,P)$ and let $\varphi_h(\cdot)$ be the kernel function of argument x. Let $x_i^{(s)}(\theta)$ be the elements of the simulated sample of model response. The density at point x has the standard expression $\widehat{f}_\theta^{(s)}(D,\,W,\,P) = \frac{1}{n}\sum_{i=1}^{n}\varphi_h(x - x_i^{(s)}(\theta))$. It is appropriate to define the derivative of the density with respect to the structural parameter as[1]:

$$\frac{\partial\,\widehat{f}_\theta}{\partial\,\theta} = \lim_{\delta\to 0}\frac{\widehat{f}_{\theta+\delta}^{(s)}(D,\,W,\,P)-\widehat{f}_\theta^{(s)}(D,\,W,\,P)}{\delta}$$

Consider the limiting behavior of this sum. First, take the expectation of the indi-

[1] The major problem here is that the size of the response sample will always coincide with the size of the dataset because the simulations are performed using the conditioning on the observation-specific covariates. This suggests that the specified limit does not necessarily exist for finite sample sizes or sequence of sample sizes approaching infinity. For the practical purposes then it is reasonable to suggest some finite δ and consider the finite difference as an approximation of the function derivative.

vidual term assuming that the density function is three times differentiable:

$$E\left\{\frac{\varphi_h(x-x_i^{(s)}(\theta+\delta))-\varphi_h(x-x_i^{(s)}(\theta))}{\delta}\right\}=\frac{\partial f_\theta(x)}{\partial\theta}$$
$$+\frac{1}{2}\frac{\partial^2 f_\theta(x)}{\partial\theta^2}\delta+\frac{1}{2}\frac{\partial^3 f}{\partial x^2\,\partial\theta}h^2\,\delta+o(\|\delta h^2\|).$$

This expression suggests that the bias due to discretization of the derivative will be of order δ. Note again that discretization error is irreducible because the derivative needs to be computed for the samples of the finite size. Therefore, such error will be present even if the sample size approaches infinity. The variance of the individual term can be obtained as:

$$V\left(\frac{\varphi_h(x-x_i^{(s)}(\theta+\delta))-\varphi_h(x-x_i^{(s)}(\theta))}{\delta}\right)=\frac{2}{\delta^2}f_\theta(x)\int h\,\varphi_h^2\,(u)\;du-$$
$$-\frac{2}{\delta^2}\left[f_\theta(x)\int\varphi_h\,(u)\;du\right]^2+o_h(\|\delta^{-2}\|).$$

As the standard CLT applies under the independence assumption, given that $n\,h_D\,h_w\,h_p\rightarrow\infty$ while $h_D\,h_w\,h_p\rightarrow 0$:

$$\sqrt{n\,h_D\,h_w\,h_p}\left(\frac{\partial\,\widehat{f}_\theta}{\partial\,\theta}-\frac{\partial f_\theta(x)}{\partial\theta}-\frac{1}{2}\frac{\partial^2 f_\theta(x)}{\partial\theta^2}\delta\right)\xrightarrow{d}N\left(0,\;\frac{2}{\delta^2}f_\theta(x)\left[\int_{-\infty}^{+\infty}\kappa^2(u)\,du\right]^3\right).$$

This expression illustrates that the attempts to reduce the bias due to discretization lead to the significant increase in the variance of the derivative estimate. A similar result hods when instead of kernel estimation one uses sieves to compute the density function. In that case the bias and the variance of density derivative with respect to the structural parameters will be inversely related.

The main conclusion from these derivation is that bruit-force gradient method is likely to fail in my case if one uses it to find the structural parameters. This also means that the expansions for the sample loss functions similar to those used to derive the asymptotic properties of M-estimators cannot be obtained for the simulation-based estimation proposed in this paper. A possible solution to this problem is to consider differentiability in $\mathbf{L}^2$ sense which will allow one to obtain correct asymptotic expansions under much lighter conditions.

B.2 Hellinger distance and asymptotic estimates

Assuming that the structural parameters of the model are contained in some compact set $\Theta \subset \mathbb{R}^k$ we can assure the convergence rate $\delta(n)$ for the density estimate of (D, W, Z) pointwise for each $\theta \in \Theta$. The reason for using the concept of Hellinger distance is that it allows one to provide the structure of the asymptotic behavior of the econometric model even in the cases where the density function approaches zero and so log-likelihood of the model approaches infinity. To extend the class of the analyzed functions even further the authors consider the concept of Hellinger differentiability which requires the square root of the density function to be differentiable in $\mathbf{L}^2$ norm. Note that under this definition "problematic" density functions such as those similar to the Epanechnikov density are differentiable even though they do not have a pointwise derivative on the entire support. Such concept can be useful in the semiparametric procedures because it allows one to avoid the problems with non-differential density functions when, for instance, B-splines of low order are used for estimation.

For probability measures $\mathbb{Q}$ and $\mathbb{P}$ the squared Hellinger distance has the expression:

$$H^2(\mathbb{P}, \mathbb{Q}) = \frac{1}{2} \int \left(\sqrt{d\mathbb{P}} - \sqrt{d\mathbb{Q}}\right)^2$$

If the likelihood ratio is defined $\mathcal{L} = \frac{d\mathbb{Q}}{d\mathbb{P}}$ then the Hellinger distance can be rewritten as:

$$H^2(\mathbb{P},\, \mathbb{Q}) = 1 - E_{\mathbb{P}}\left\{\sqrt{\mathcal{L}}\right\}.$$

Let us adopt this definition for my case so that let $\widehat{f}_{\theta_0}(\cdot)$ be the density of the data and $\widehat{f}_{\theta}^{(s)}(\cdot)$ be the density of the simulated response. For notational simplicity denote $x = (D,\, W,\, Z)$. Then for the empirical density measures we can write the expression for the Hellinger distance as:

$$\widehat{H}_n(\theta,\, \theta_0) = 1 - \frac{1}{n}\sum_{i=1}^{n}\sqrt{\frac{\widehat{f}_{\theta}^{(s)}(x_i)}{\widehat{f}_{\theta_0}(x_i)}}$$

Consider the Hellinger expansion of the distance for $\xi_i(\theta,\, \theta') = \sqrt{\widehat{f}_{\theta}^{(s)}(x_i)/\widehat{f}_{\theta'}(x_i)}$. In this case the density expansion in $\mathbf{L}^2$ sense can be determined as:

$$\sqrt{f_\theta(x)} = \sqrt{f_{\theta_0}(x)} + \frac{1}{2}(\theta - \theta_0)'\Delta(x)\,\sqrt{f_{\theta_0}(x)} + r_\theta(x).$$

Note that $\widehat{f}_\theta(x)$ is differentiable in $\mathbf{L}^2$ sense and the same expansion can be written for it. Therefore, the Hellinger distance can be represented as:

$$\widehat{H}_n(\theta,\, \theta_0) = 1 - \\ -\frac{1}{n}\sum_{i=1}^{n}\frac{\sqrt{\widehat{f}_{\theta_0}^{(s)}(x_i)} + \frac{1}{2}(\theta-\theta_0)'\Delta_1^{(s)}(x_i)\,\sqrt{\widehat{f}_{\theta_0}^{(s)}(x)} + \frac{1}{4}(\theta-\theta_0)'\Delta_2^{(s)}(x_i)(\theta-\theta_0)\,\sqrt{\widehat{f}_{\theta_0}^{(s)}(x)} + r_\theta(x)}{\sqrt{\widehat{f}_{\theta_0}(x_i)}}.$$

As $\widehat{f}_\theta(\cdot)$ is asymptotically normal, we can find a random variable $\xi(x)$ and a parameter $\eta(n) = (n\, h_d\, h_w\, h_z)^{-1}$ so that $E\left[\xi^2(x)\right] < M$ for some positive constant M and $\widehat{f}_{\theta_0}^{(s)}(x) = f_{\theta_0}(x) + \xi(x)\,\eta(n)$, and $\widehat{f}_{\theta_0}(x) = f_{\theta_0}(x) + \widetilde{\xi}(x)\,\eta(n)$. Substituting these expressions into the expression for the Hellinger distance:

$$\widehat{H}_n(\theta,\, \theta_0) = \frac{1}{2}(\theta - \theta_0)' \frac{1}{n} \sum_{i=1}^{n} \Delta_1^{(s)}(x_i) + \frac{1}{4}(\theta - \theta_0)' \frac{1}{n} \sum_{i=1}^{n} \Delta_2^{(s)}(x_i)(\theta - \theta_0) + \frac{\Lambda_n\, \eta(n)}{2},$$

where $E\left[\Lambda_n^2\right] < \infty$. Now we can analyze this transformed distance function to obtain the asymptotic behavior of the minimum distance estimates. For this purpose we can use the results from the theory of empirical processes, for instance, from [And94b].

Consider a single realization of the stochastic process $\{x_t^{(k)}\}_{t\in[0,T]}$ $= \{D_t^{(k)}, W_t^{(k)}, Z_t^{(k)}\}_{t\in[0,T]}$. For this realization one can write:

$$\sum_{i=1}^{N_\tau} \sqrt{\left[\frac{\widehat{f}(x_i^{(k)})}{\widehat{f}_\theta(x_i^{(k)})}\right]} = \int_0^T \sqrt{\left[\frac{\widehat{f}(x_t^{(k)})}{\widehat{f}_\theta(x_t^{(k)})}\right]}\, dP_{N_\tau},$$

where $P_{N_\tau}(\cdot)$ is the empirical measure, generating the discretized stochastic process $\{x_t^{(k)}\}_{t\in[0,T]}$. Let ξ_1 and ξ_2 are nuisance parameters in non-parametric estimates of $f_{\theta_0}(\cdot)$ and $f_\theta(\cdot)$. Let us denote the latter stochastic integral by $J^{(k)}\left(\theta,\, \xi_1,\, \xi_2\right)$.

Note now that the Hellinger distance can be written as

$$\widehat{H}\left(\theta_0,\, \theta,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) = 1 - \frac{1}{n} \sum_{k=1}^{n} J^{(k)}\left(\theta,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) = 1 - J_n\left(\widehat{\theta},\, \widehat{\xi}_1,\, \widehat{\xi}_2\right).$$

Note that the structure of the estimator is similar to that of M-estimators, but the deterministic moment condition in this case is substituted by a stochastic moment condition. Under correct specification the estimator for θ_0 satisfies:

$$J_n\left(\widehat{\theta},\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) - 1 = 0$$

with probability 1 as $n \to \infty$.

To prove the consistency and derive asymptotic properties of the obtained estimator we follow first the reasoning in [And94b]. In fact note that one can write down the mean value expansion as:

$$o_p(1) = J_n\left(\widehat{\theta},\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) - 1 = \sqrt{n}J_n\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) + \frac{\partial}{\partial\theta'}J_n\left(\theta^*,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right)\sqrt{n}\left(\widehat{\theta} - \theta_0\right) - 1$$

As the integral over empirical measure converges to the integral over the Wiener measure:

$$\frac{\partial}{\partial\theta}J_n\left(\theta^*,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) \xrightarrow{p} -Q = -E\left\{\int_0^T \frac{\partial}{\partial\theta}\sqrt{f\left(x_t\theta_0\right)}\, dP_t\right\}$$

The latter integral is finite as long as the distribution has a mean square expansion of its density.

The expression for the parameter under consideration can be written as:

$$\begin{aligned}&\sqrt{n}\left(\widehat{\theta} - \theta_0\right) = Q^{-1}\sqrt{n}J_n\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) + o_p(1) = \\ &= Q^{-1}\left[\sqrt{n}\left(J_n\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) - J_n^*\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right)\right) + \sqrt{n}J_n^*\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right)\right],\end{aligned} \tag{B.1}$$

where $J_n^*\left(\theta,\, \xi_1,\, \xi_2\right) = \frac{1}{n}\sum\limits_{k=1}^{n} E\{J^{(k)}\left(\theta,\, \xi_1,\, \xi_2\right)\}$.

As ξ_1 and ξ_2 are nuisanse parameters in the non-parametric density estimation, it has been shown in above that, similarly to standard kernel estimators as in [Sil86] such estimates are pointwise asymptotically normal. Using Fubbini theorem one can argue that

$$var_\xi\left[J^{(k)}\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right)\right] = E\{\textstyle\int_0^T var_\xi\left[\sqrt{\left[\frac{\widehat{f}(y_t,t)}{\widehat{f}_{\theta_0}(y_t,t)}\right]}\right]\, dP_t\}.$$

The existence of this variance is justified by the existence of finite variances of $\sqrt{\left(\widehat{f}(\cdot)\right)}$ and $\sqrt{\left(\widehat{f}_{\theta}(\cdot)\right)}$ which is guaranteed by the existence of the mean square expansion for the square root of the density. In particular, if the simulated sample is independent from the actual sample of trajectories we can write the asymptotic expression for the variance given that the density estimates are obtained from the kernel smoother with a kernel function $\kappa(\cdot)$ given assumptions A4 and A5 as:

$$var_{\xi}\left[\sqrt{nh_t h_y} J_n^*\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right)\right] \stackrel{n\to\infty}{\longrightarrow} \left(\int_0^{\infty} \kappa^2(\psi)\, d\psi\right)^2 E\left\{\int_0^T \frac{dP_t}{\sqrt{f(y_\tau, \tau, \theta_0)}}\right\} = \Omega_{\xi}$$

Note that this variance is only driven by the variance in the estimation of joint density but not by the stochastic process process per se. The reason for this is that if we could have a perfect estimate of the distribution of the process x and the timing of its jumps, then the Hellinger distance under true parameter values will be equal to zero and thus the variance of the corresponding stochastic integral would be equal to zero as well. The only source of variance in J_n^* is therefore the error in non-parametric density estimate.

This integral exists as the considered Wiener measure has a Radon-Nykodim density.

We can further write then that:

$$\sqrt{nh_t h_y} J_n^*\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) \xrightarrow{d} N\left(0, \Omega_{\xi}\right).$$

Denote now $v_n(\xi) = \sqrt{n}\left(J_n\left(\theta_0,\, \xi_1,\, \xi_2\right) - J_n^*\left(\theta_0,\, \xi_1,\, \xi_2\right)\right)$. Under true $\xi = \xi^0$ as $J^{(k)}(\cdot)$ are independent while $v_n(\xi^0) \equiv 0$.

Assuming that the mean square expansion of the density can be evaluated with the Lipschitz constant Λ, then:

$$E\left\{\sup_{\|\xi-\xi'\|\leq\delta}\left|J^{(k)}\left(\theta_0,\xi'\right)-J^{(k)}\left(\theta_0,\xi\right)\right|^2\right\}\leq 4\Lambda\delta^2,$$

by Doob's inequality. Then given mean - square convergence of non-parametric estimates of density under $\mathbf{L}^2$ norm, according to [And94b], $v_n(\cdot)$ is stochastically equicontinuous in ξ which implies that $v_n\left(\widehat{\xi}\right)-v_n\left(\xi^0\right)\xrightarrow{p}0$.

This assures that the estimate $\widehat{\theta}$ minimizing the Hellinger distance is asymptotically normal and:

$$\frac{1}{\sqrt{\eta(n)}}\left(\widehat{\theta}-\theta_0\right)\xrightarrow{d}N(0,\,Q^{-1}\,\Omega_\xi\,Q^{-1}),$$

where $Q=\frac{\partial}{\partial\theta}E\left\{\sqrt{\widehat{f}_{\theta_0}(x_i)}\right\}$ and $\Omega_\xi=\frac{\partial}{\partial\theta\,\partial\theta'}E\left\{\sqrt{\widehat{f}_{\theta_0}(x_i)}\right\}$.

Note that to make this expansion one does not need differentiability of the estimated density with respect to the structural parameters. The imposed requirement concerns the differentiability of the expected estimate of the density with respect to the structural parameter. The existence of this derivative is ensured by the existence of the derivative of the true density function $f_\theta(\cdot)$ with respect to parameter θ which seems to be a reasonable requirement. Moreover, even this assumption can be weakened by only requiring that the integral $\int\left\{f_\theta(x)\right\}^{3/2}\,dx$ is differentiable with respect to θ. This assumption is likely to hold even in the non-regular cases when the support of the distribution $f_\theta(\cdot)$ depends on the parameter or when the density has first-order discontinuities.

Table B.1: General Data for Executive Managers

variable	N. obs.	mean	stdev	min	max
Salary (thous. $)	1851	430.7159	305.2276	.001	2279.138
Bonus (thous. $)	1851	268.1957	500.8492	0	5672.5
Executive age	651	54.4424	7.5858	42	78
Gender (Female=1)	1851	.2106969	.4079137	0	1

Table B.2: Some Characteristics of Compensations (thousands $)

variable (values in thous. $)	N. obs.	mean	stdev	min	max
Total compensation	1851	698.8374	733.2091	.001	7826.584
Restricted stock grants	1851	235.2837	1504.514	0	40250
Restricted stock holdings (value)	1630	44.65941	258.5092	0	4500
Options granted (B-S value)	1851	122.6672	422.6474	0	5000
Long term incentive payouts	1851	17.35361	146.881	0	3484.613
Options exercised	1630	654.7266	3015.383	0	88491.23
Unexercised exercisable op.	1630	233.4002	900.3827	0	14523.19
Unexercised unexercisable op.	1630	373.8253	1322.936	0	19629.75
Tot. compens. change (% per year)	1501	454.6537	15378.86	-92.954	595551.3

Table B.3: Structural estimates of demand and contract parameters

Parameter	Estimate	Std. deviation
	Drift in demand	
a_0^ψ	2.9296	1.6103 **
a_1^ψ	-0.1407	0.7875
a_2^ψ	-0.0310	1.4101
a_3^ψ	-2.9764	0.6215 ***
	Diffusion in demand	
a_0^ζ	0.5304	0.3612
a_1^ζ	-1.8909	0.6208 ***
a_2^ζ	-0.2784	0.9215
a_3^ζ	3.5345	1.6106 **
	Diffusion in tastes	
a_0^γ	-1.5086	0.6848 **
a_1^γ	-2.4000	1.2752 **
a_2^γ	1.553593	0.9050 **
a_3^γ	1.5478	1.1609 *
	Cost function	
a_0^c	1.3012	0.5834 **
a_1^c	0.1607	0.0830 **
a_2^c	-1.2324	1.1545
	Reward for level of sales	
a_0^α	0.3655	0.7095
a_1^α	-0.0092	0.5886
a_2^α	-0.6352	0.9920
a_3^α	-0.1107	0.4291
	Reward for growth of sales	
a_0^β	-0.3333	0.6646
a_1^β	0.7126	0.3604 **
a_2^β	-0.3386	0.5106
a_3^β	-1.5050	0.4819 ***
	Wage and risk aversion	
σ	0.9577366	0.4618519 **
R_m	0.8246784	0.4039156 **

(***) reflects significance on 1% level, (**) - 5% significance, (*) - 10% significance

Table B.4: Structural estimates of parameters of the single index

Variable	Estimate	Std. deviation
Sales % change	-0.1380	0.0977
Market value	-1.1976	0.5734 **
Pre-tax/average income	0.7703	0.5608
Book-to-market ratio	0.1993	0.4151
Stock grant options/options	1.6036	0.9515
Option grants	2.0901	0.5417 ***
Value of shares/bonus	-3.3760	1.5660 **
Gender (female=1)	-2.5012	1.5335 *

Table B.5: Estimation results for the system describing firm's sales and manager's compensation

Variable	Coefficient	std. error	t-statistic
Dependent variable: Salary ($\mathbf{R}^2$=0.212)			
Sales	51.51886	2.50402	20.57
Firm's value	0.0021807	0.0004241	5.14
Tobin's q	2.452688	0.5870775	4.18
Change in assets	-0.0034975	0.0035935	-0.97
Apparel	67.90431	7.229979	9.39
SOX	27.62433	11.58859	2.38
Time	14.29167	0.9717439	14.71
Constant	292.5628	6.395592	45.74
Dependent variable: Bonus ($\mathbf{R}^2$=0.189)			
Sales	70.91936	4.438576	15.98
Firm's value	0.0090904	0.0007518	12.09
Tobin's q	11.67322	1.041005	11.21
Change in assets	-0.0048182	0.0063721	-0.76
Apparel	51.95567	12.8201	4.05
SOX	43.49443	20.54893	2.12
Time	19.77868	1.723088	11.48
Constant	74.17933	11.34068	6.54
Dependent variable: Sales ($\mathbf{R}^2$=0.414)			
Firm's value	0.0001162	1.76E-06	66.01
Tobin's q	-0.0233784	0.0033326	-7.02
Change in assets	-0.0000402	0.0000205	-1.96
Apparel	-0.4840632	0.0406628	-11.9
SOX	-0.034501	0.0661104	-0.52
Time	-0.0556195	0.0054865	-10.14
Constant	0.1261532	0.036442	3.46
N. obs =6756			

Appendix C
Proofs and Figures in Chapter 3

C.1 The structure of the price process

In this appendix I provide conditions for timing of individual bids. Suppose that the bidder controls both the bid and its timing. The bid is a piecewise continuous function u_t and du_t is a a linear combination of Dirac's delta functions indicating instants when bids are submitted. Take function $f_\psi(\cdot,\,\cdot)$ to be a smoothed step function parametrized by a single parameter ψ, such that $f_\psi(x,\,y) \xrightarrow{\psi\to 0} \max\{x,\,y\}$. Denote by x_t the path of the state variable when the considered bidders is removed and y_t is the price path taking into account bids of the considered bidder. The increments of the state variable process y_t can be defined by the increments of the process x_t and the increments of control. Therefore as $d\,y_t = \max\{dx_t,\,du_t\}$ we can define a process $\{y_t^\psi\}$ with increments $d\,y_t^\psi = df_\psi(x_t,\,u_t)$. According to [SM91] as x_t is pure jump process we can represent this expression as

$$d\,y_t^\psi = (f_\psi(x_t + h(x_t,\,t),\,u_{t+dt}) - f_\psi(x_t,\,u_{t+dt}))dJ(t,\,x_t) + (f_\psi(x_t,\,u_{t+dt}) - f_\psi(x_t,\,u_t)).$$

This expansion reflects the equality in the mean square.

We can always chose a twice continuously differentiable concave increasing function $f_\psi(\cdot,\,\cdot)$ which means that Lipschitz conditions will be valid for its derivatives. Therefore, first, the process y_t^ψ can be written in a unique way as:

$$\begin{aligned} y_t^\psi = \int_0^t (f_\psi(x_t + h(x_t,\,t),\,u_{t+dt}) - f_\psi(x_t,\,u_{t+dt}))dJ(t,\,x_t) \\ + \int_0^t (f_\psi(x_t,\,u_{t+dt}) - f_\psi(x_t,\,u_t)) \end{aligned} \tag{C.1}$$

Due to [GS79], as the transition function is continuous and converges to max - function, then $y^{\psi}_t \to y_t$ weakly. As $f_\psi(\cdot, \cdot)$ is increasing then $\int\limits_0^t (f_\psi(x_t, u_{t+dt}) - f_\psi(x_t, u_t)) \geq 0$. Due to concavity of function f_ψ $u_{t+dt} > u_t$ can only be optimal for the minimum of $y^{\psi})_t$ if $dJ(t, x_t) > 0$. This means that the control is applied only in the moments of jumps.

C.2 Uniqueness of Individual Best Responses

Consider an operator Γ such that for a bounded function $g(x, y)$ defined on $[0, T]\times [0, \overline{Y}]$ we have $\Gamma \circ g(x, y) = x\, y/T - \int\limits_x^T \lambda(t, y)\, (g(t, y + h(t, y)) - g(t, y))\, dt$. Suppose that functions $\lambda(x, y)$ and $h(x, y)$ are continuous and bounded on $[0, T] \times [0, \overline{Y}]$.

Let $\|u(x, y) - y\| < b$ for $y \in [0, \overline{Y}]$ and $|x - T| < a$ for some positive constants a and b. Let these conditions hold for some functions $u_1(x, y)$ and $u_2(x, y)$.

As the functions forming the operator are continuous and bounded, we can find a constant c determined by their Lipschitz constants[1] such that

$$\|\Gamma\, u_1(x, y) - \Gamma\, u_2(x, y)\| \leq c\|u_1 - u_2\|\, \|x - T\|$$

The same procedure can be applied to the operator Γ^2 (as Γ can be applied to the function $\Gamma \circ g(x, y)$) which gives the expression:

$$\|\Gamma^2\, u_1(x, y) - \Gamma^2\, u_2(x, y)\| \leq c^2\|u_1 - u_2\|\, \|x - T\|^2/2$$

[1]If we impose the restriction that $y + h(x, y) \in [0, \overline{Y}]$ for all $x \in [0, T]$, then we will only need that both functions λ and h are bounded to have this result. Lipschitz property will be required to perform consequent steps.

Iterating this procedure we obtain for Γ^p:

$$\|\Gamma^p u_1(x,\, y) - \Gamma^p u_2(x,\, y)\| \leq c^p \|u_1 - u_2\|\, \|x - T\|^p/(p!)$$

Choosing big enough p we will have that $\frac{c^p \, \sup \|x-T\|^p}{p!} < \alpha < 1$ and

$$\|\Gamma^p u_1(x,\, y) - \Gamma^p u_2(x,\, y)\| \leq \alpha \|u_1 - u_2\|$$

Therefore Γ^p is a contraction mapping for the considered class of bounded functions. By [DS58] the operator Γ has a unique fixed point if Γ^p is a contraction mapping for some $p \in \mathbb{N}$. This implies that the equation defined as $g = \Gamma \circ g$ has a fixed point which is unique. Such fixed point is the solution to the equation (3.7). This proves the necessary result.

C.3 Derivation of the Optimal Filter

The derivation of the optimal filter here will follow that in the paper [VS77], where the expression for the general filtering of count processes was derived.

Let us look for the solution to the filtering problem to be of the form

$$d\theta_t = \xi_t \frac{1}{h(t, x_t, \theta_t)} \left\{ dx_t - \psi(t,\, x_t,\, \theta_t)\, dt \right\}.$$

Denote $dJ(\theta) = dx_t - \psi(t,\, x_t,\, \theta_t)\, dt$. Therefore one can write:

$$dx_t = \psi(t,\, x_t,\, \theta^*)dt + h(t,\, x_t,\, \theta^*)dJ(\theta^*),$$

where θ^* is the true parameter value. Denote $e_t = \theta_t - \theta^*$ and $e_t^{\psi} = \psi(t,\, x_t,\, \theta_t) -$

$\psi(t,\, x_t,\, \theta^*),\ \psi^* = \psi(t,\, x_t,\, \theta^*),\ \psi_t = \psi(t,\, x_t,\, \theta_t),\ h^* = h(t,\, x_t,\, \theta^*),\ h_t = h(t,\, x_t,\, \theta_t)$. We will also use the notation $[\cdot,\, \cdot]_t$ for quadratic covariation between the processes.

Using the introduced notations we can write:

$$dJ(\theta_t) = -\frac{1}{h_t} e_t^{\psi} + \frac{h^*}{h_t} dJ(\theta^*)$$

Using the rules of stochastic calculus for count processes we can write the following sequence of calculations:

$$\begin{aligned} & e_t J(\theta_t) = e_s J(\theta_s) + \int_s^t e_{\tau-}\, dJ(\theta_\tau) + \int_s^t J(\theta_{\tau-})\, de_\tau + \int_s^t d[e,\, J(\theta_t)]_t = \\ & = e_s J(\theta_s) - \int_s^t \tfrac{e_{\tau-}\, e_\tau^{\psi}}{h_\tau}\, d\tau + \int_s^t \tfrac{h^*}{h_\tau} e_\tau\, dJ(\theta^*) + \int_s^t J(\theta_\tau)\xi_\tau\, dJ(\theta_\tau) + \int_s^t \xi_\tau \left\{ \tfrac{\psi_\tau}{h_\tau}\, d\tau + dJ(\theta_\tau) \right\} \end{aligned}$$

Due to [VS77] $E\{e_t\, J(\theta_t) | \mathfrak{F}_{ns}\} = 0$ for sample-generated σ-algebras $\mathfrak{F}_{ns}$. Also one can conclude that $E\{e_t\, J(\theta_t) | \mathfrak{F}_{ns}\} = E\{\int_s^t \left\{ -\frac{e_\tau e_\tau^{\psi}}{h_\tau} + \xi_\tau \frac{\psi_\tau}{h_\tau} \right\} d\tau | \mathfrak{F}_{ns}\}$.

From this it follows that $\xi_t = \frac{E\{e_\tau e_\tau^{\psi} | \mathfrak{F}_{ns}\}}{E\{\psi_\tau | \mathfrak{F}_{ns}\}}$.

Note that in the first approximation $e_t^{\psi} = \frac{\partial \psi(t, x_t, \theta_t)}{\partial t}(\theta_t - \theta^*)$, therefore we get the following formula for the linear approximation of the filtered process: $\xi_t = \frac{\partial \psi_t}{\partial \theta} \frac{\sigma_t(\theta_t)}{\psi_t}$, where $\sigma_t(\theta_t) = E\{(\theta_t - \theta^*)^2 | \mathfrak{F}_{ns}\}$.

Now this defines the dynamics of the conditional variance of the estimate given the third moment σ_3. Consider a stochastic differential for the function $\psi(\cdot)$:

$$d\psi(t, x_t, \theta_t) = \left\{ \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial x}\psi - \frac{\partial \psi}{\partial \theta} \frac{\xi_t}{h_t} e^{\psi} \right\} dt + \{\psi(t, x_t + h_t, \theta_t + \frac{h^*}{h_t}\xi_t) - \psi(t, x_t, \theta_t)\}\, dJ(\theta^*)$$

Denote $A = \frac{\partial}{\partial \theta}\left\{ \frac{\partial \psi}{\partial t} + \frac{\partial \psi}{\partial x}\psi \right\}$ and $B = \frac{\partial \psi(t, x_t + h_t, \theta_t}{\partial \theta} - \frac{\partial \psi(t, x_t, \theta_t)}{\partial \theta}$.

Write:

$$de_t^{\psi} = \left\{ A\, e_t - \frac{\partial \psi}{\partial \theta} \frac{\xi_t}{h_t} e_t^{\psi} \right\} dt + B\, e_t\, dJ(\theta^*).$$

Then for the differential of the quadratic covariation:

$$d[e,\, e^{\psi}] = B\, e_t \xi_t \frac{\psi_t}{h_t}\, dt + B\, e_t \xi_t\, dJ(\theta^*).$$

Therefore:

$$d(e_t\, e_t^{\psi}) = e_{t-} de_t^{\psi} + e_{t-}^{\psi} de_t + d[e,\, e^{\psi}] = \left\{ e_t (A\, e_t - \tfrac{\partial \psi}{\partial \theta} \tfrac{\xi_t}{h_t} e_t^{\psi}) - e_t^{\psi} \xi_t \tfrac{e^{\psi}}{h_t} + B\, e_t \xi_t \tfrac{\psi_t}{h_t} \right\} dt +$$
$$+ r_t^e\, dJ(\theta^*) = \left\{ A\, e_t^2 - \tfrac{\partial \psi}{\partial \theta} \xi_t \tfrac{e_t^{\psi} e_t}{h_t} - \xi_t \tfrac{(e_t^{\psi})^2}{h_t} + B\, e_t \xi_t \tfrac{\psi_t}{h_t} \right\} dt + r_t^e\, dJ(\theta^*)$$

Applying the same tools as for deriving the main equation, for $r_t = E\{r_t^e | \mathfrak{F}_{ns}\}$ the expression in terms of expectations looks like $r_t = \frac{E\{e_\tau (e_\tau^{\psi})^2 | \mathfrak{F}_{ns}\}}{E\{\psi_\tau / h_\tau | \mathfrak{F}_{ns}\}}$.

Using the first-order expansion this expression transforms to: $r_t = \left(\frac{\partial \psi}{\partial \theta}\right)^2 \frac{h_t\, \sigma_3}{\psi_t}$.

Therefore, the equation for the dynamics of $\sigma_t(\theta_t)$ can be calculated by using the expression for r_t and conditioning of the drift for the process $e_t\, e_t^{\psi}$ and using the expansion for e^{ψ}:

$$d\sigma_t(\theta_t) = \left\{ A \left(\frac{\partial \psi}{\partial \theta} \right)^{-1} - \frac{1}{h_t \psi_t} \left(\frac{\partial \psi}{\partial \theta} \right)^2 \right\} \sigma_t^2(\theta_t)\, dt + \frac{\partial \psi}{\partial \theta} \frac{\sigma_3}{\psi_t} \left\{ dx_t - \psi(t,\, x_t,\, \theta_t)\, dt \right\}.$$

By letting $\sigma_3 \equiv 0$, this concludes the derivation of (3.5).

C.4 Numerical Example

The problem (3.7) can be solved by standard methods used to solve partial differential equations. This partial differential equation, however, contains a significant non-linearity introduced by the supremum operator. This does not allow one to use efficient computational techniques such as the method of fractional steps (which allows one to reduce the size of the computation grid without a large effect on the precision of integration) to directly solve it. It is well known that for partial differential equations of the diffusion type, explicit Euler's forward integration method produces an unstable numerical solution. For this reason, it is necessary to use a computational procedure that allows one to exploit the benefits of the fractional step method, but possesses the stability property of the implicit integration scheme. One such method is the fractional - steps method with Crank-Nicholson correction step. I present this algorithm for a particular example. Consider the optimal bidding behavior when price jumps are described by $h(t, x, a_t) = \max\{h(t, x), a_t\}$.

Example. I first present the algorithm and then discuss its speed and stability. As in many other boundary problems, this problem is solved backwards, using boundary condition as a starting point. Let $Z_{t\,i}$ be an array with the values of Z with index $t = 1, \ldots, T$ referring to time labels and index $i = 1, \ldots, N$ referring to the state variable labels. Let set $I_{t\,i}$ refer to the set of $\inf_{a_t > 0} Z(t, x + \max\{h(x,\, t),\, a_t\})$ for each time label t and state variable label i. Let the points x_i form the grid for the state variable and let τ be the time step.

Step 0.

Initialize $Z_{T\,i} = x_i$. Initialize $I_{T\,i} = x_i + h(x_i,\, T)$.

Step 1.

Extrapolate the value function at an intermediate point. In the beginning of integration extrapolation is equal to the boundary value. After two steps extrapolate the value function as:

$$Z^{ex}_{t-0.5\,i} = 0.5\,(3Z_{i,\,t} - Z_{i\,t+1})$$

Step 2.

Compute corrected value function using the extrapolated value to compute spatial derivatives and finite differences, i.e. for every i

$$Z^{(c,1)}_{t-1\,i} = Z^{ex}_{t\,i} + \tau\,\lambda_\epsilon(x_i,\,t)\,(I_{t\,i} - Z^{ex}_{t\,i})$$

Step 3.

Update the value function in the intermediate point as:

$$Z^{(c,1)}_{t-0.5,i} = 0.5\left(Z^{(c,1)}_{t+1,i} + Z_{t,i}\right).$$

Repeat Step 2 with $Z^{(c,1)}_{t-0.5,i}$ for spatial derivatives and finite differences to get $Z^{(c,2)}_{t-1,i}$. Repeat Steps 2-3 until convergence is achieved or the number of iteration steps exceeds the maximum (in our experiment we set the maximum number of iterations to 50 which has never been achieved) to get $Z^{(c,k_{max})}_{t-1,i}$.

Step 3.

Set $Z_{t-1,i} = Z^{(c,k_{max})}_{t-1,i}$

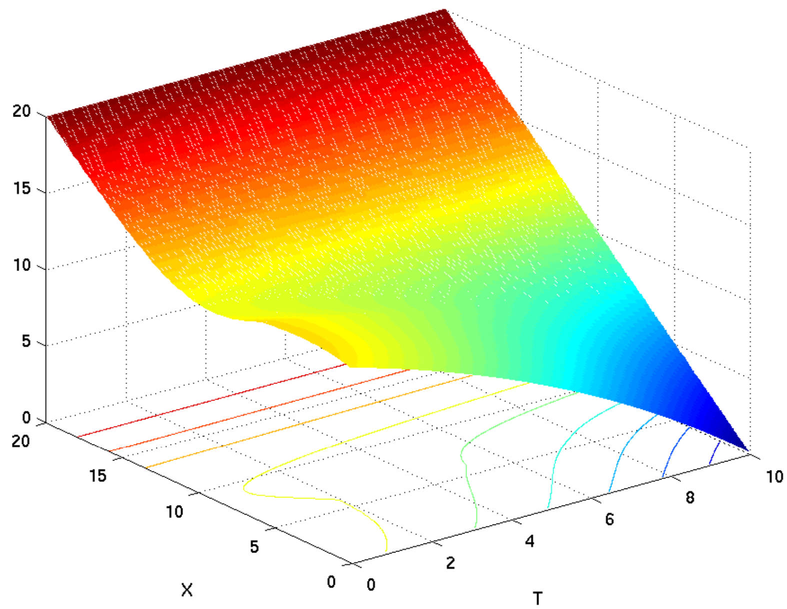
Step 4.

Find $\min_{j\geq i}\{Z_{t-1\,j}\}$ for every i. If $argmin_{j\geq i}\{Z_{t-1\,j}\} \neq i$, then $I_{t-1\,i} = Z_{t-1\,i}$ otherwise take the grid point x_j closest to the point $x_i + h(x_i)$ and set $I_{t-1\,i} = Z_{t-1\,i}$.

Steps 1- 4 should be repeated until $t = 0$.

From the practical viewpoint, an important problem in implementing the algo-

Figure C.1: Computation result for the time-stationary problem.

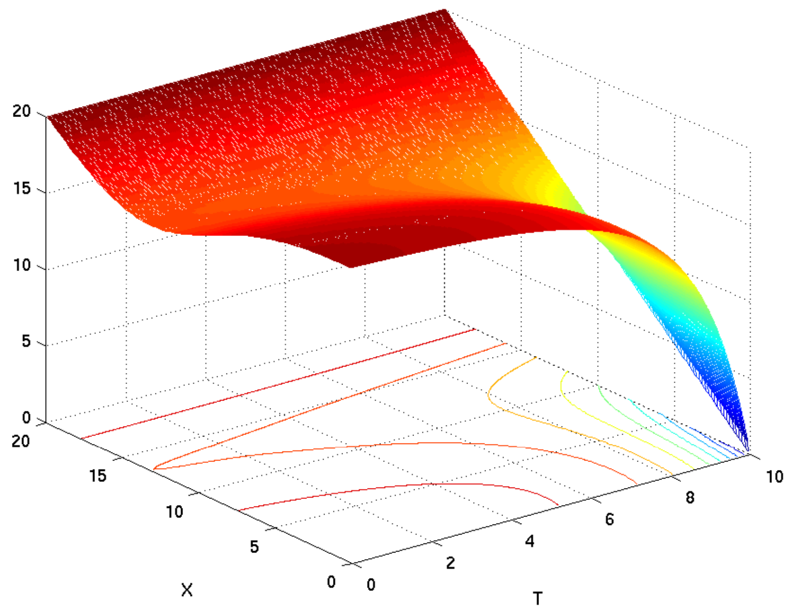


rithm to use for real data is calibration of parameters of Poisson process. If the frequency of price jumps does not decay fast enough if the state variable increases, the numerical solution to the analyzed equation can become explosive. Intuitively this restriction means that the frequency of jumps in the state variable should be predictable by the players.

I illustrate the influence of time on the expected final value of state variable with the following functions. First I consider a time-stationary problem with the functions $\lambda(x,\, t) = \alpha_1 e^{-\beta_1 x^2}$ and $h(x,\, t) = \alpha_2 e^{-\beta_2 x^2}$. The parameters of the functions were $\beta_1 = \beta_2 = 0.01$, $\alpha_2 = 20$ and $\alpha_1 = 0.1$. Figure (C.1) below shows the surface corresponding to the solution of the boundary problem (3.7) with these parameters.

The graph shows the values of the expected final price given current price for every moment of sale duration under the optimal control from the buyer. The picture is consistent with the simulation results. In fact for the moments of time close to the

Figure C.2: Model with exponentially growing parameters



beginning of the auction, the expectation for the final price given optimal control is a function with a set minima for a region of relatively small initial prices, while for higher prices there is a one-to-one correspondence between the initial price and the expected current price. Similarly, as predicted by the simulations, the dependence tends towards a one-to-one correspondence for most of the price levels when the sale is close to the end.

Adding a time dimension to the problem changes the structure of solution. The characteristics of the Poisson process now look like $\lambda(x,\, t) \;=\; \alpha_1 e^{-\beta_1\, x^2}\, e^{\gamma_1\, t}$ and $h(x,\, t) \;=\; \alpha_2 e^{-\beta_2\, x^2}\, e^{\gamma_2\, t}$. The added parameters γ_1 and γ_2 were set equal to 0.1. The expected price at the end of the sale process generated under such parameter values is shown on the figure (C.2).

The graph shows that adding an exponentially growing time component to the parameters of the process inflates the expectation of the final price at the beginning of the sale. The comparison of the graphs for time-dependent and time-independent processes demonstrates that for a wide range of initial values the expected final price is higher for the time-dependent process. However, the general structure of the contour sets of the solution remains the same with a minimum and convergence to a one-to-one correspondence for high initial values of the process.

The existence and uniqueness of the solution of problem (3.7) can be proved in the same way it is proved for the the boundary problems for the partial differential equations. The proof is based on the fact that the boundary problem (3.7) can be rewritten as an integral equation. The solution to problem (3.7)is, therefore, also a solution to operator equation with the integral operator. Due to the result from operator theory [2], a solution to operator equation exists and is unique if some power of the generating operator is a contraction operator. The complete proof is given in the previous part of Appendix C. The same technique can be used to prove the convergence of the chosen algorithm. In this case the operator can be rewritten in terms of finite differences, while the result remains the same.

This example demonstrates that the problem of the agent can be solved by a simple linear algorithm that does not require value function iterations. Moreover, as the problem does not contain the derivatives of order higher than one, the requirements for the calculation grid are quite mild.

C.5 Asymptotic properties

Consistency and asymptotic normality

Given assumptions **A1** - **A5** let us show that (9) will be pointwise converging in

[2] [DS58]

probability to the marginal distribution, that is:

$$\widehat{f}(x,\, t) \overset{n\to\infty}{\longrightarrow} \int f(x,\, \delta,\, t,\, \theta_0)\, d\delta$$

For a single realization k:

$$\begin{aligned}\xi_k = \tfrac{1}{h_x\, h_t} \sum_{i=1}^{I_k} \kappa\left(\tfrac{x_i^{(k)}-x}{h_x}\right) \kappa\left(\tfrac{t_i^{(k)}-t}{h_t}\right) \\ = \tfrac{1}{h_x\, h_t} \int_0^T \kappa\left(\tfrac{x_\tau^{(k)}-x}{h_x}\right) \kappa\left(\tfrac{\tau-t}{h_t}\right) dJ(\tau, x_\tau^{(k)}, \delta_\tau^{(k)}),\end{aligned}$$

by definition where $J(\cdot)$ is the Poisson measure generating the stochastic process $\left\{x^{(k)},\, t^{(k)},\, \delta^{(k)}\right\}$. This integral has finite expectation and variance due to assumptions A1 and A2 and generalized Cauchy-Schwartz inequality.

Note also that

$$\begin{aligned}f(x,t) = \lim_{\substack{\Delta_t\to 0\\ \Delta_x\to 0}} \tfrac{1}{\Delta_x\, \Delta_t} E\left\{\int_x^{x+\Delta_x} \int_t^{t+\Delta_t} dJ(\tau, x_\tau, \delta_\tau)\right\} = \lim_{\substack{\Delta_t\to 0\\ \Delta_x\to 0}} E\left\{\eta(x,t,\Delta_x,\Delta_t)\right\} \\ = \lim_{\substack{\Delta_t\to 0\\ \Delta_x\to 0}} \tfrac{1}{\Delta_x\, \Delta_t} E\left\{\int_0^T \mathbf{1}\{|x_\tau - x| < \Delta_x\}\, \mathbf{1}\{|\tau - t| < \Delta_t\} dJ(\tau, x_\tau, \delta_\tau)\right\}.\end{aligned}$$

Then

$$\begin{aligned}E\left\{\sup|\xi(x,t) - \eta(x,t,h_x,h_t)|^2\right\} = E\left\{\sup\left|\int_0^T \left[\tfrac{1}{h_x\, h_t}\kappa\left(\tfrac{x_\tau - x}{h_x}\right)\kappa\left(\tfrac{\tau-t}{h_t}\right) - \right.\right.\right. \\ \left.\left.\left. -\tfrac{1}{h_x\, h_t}\mathbf{1}\{|x_\tau - x| < h_x\}\, \mathbf{1}\{|\tau - t| < h_t\}\right] dJ(\tau, x_\tau, \delta_\tau)\right|^2\right\} \leq \\ \leq 4\, E\left\{\int_0^T \int_0^\infty \left[\tfrac{1}{h_x\, h_t}\kappa\left(\tfrac{x_\tau - x}{h_x}\right)\kappa\left(\tfrac{\tau-t}{h_t}\right) - \tfrac{1}{h_x\, h_t}\mathbf{1}\{|x_\tau - x| < h_x\}\, \mathbf{1}\{|\tau - t| < h_t\}\right]^2\right. \\ \left. \lambda(\tau, x_\tau, \delta_\tau)\, dx_\tau\, d\tau\right\} \to 0,\end{aligned}$$

as h_t and h_x approach 0 due to Doob's inequality, Cauchy-Schwartz inequality and the fact that the kernel function approaches the step function in $\mathbf{L}^2$ norm. This implies that $E\{\xi(x,t)\} \to f(x,t)$ as h_x and $h_t \to 0$.

Due to [GS79] we can express the variance of $\xi(x,t)$ as:

$$var\left(\xi(x,t)\right) = \frac{1}{h_t^2 h_x^2} \int_0^T \int_0^\infty \kappa^2\left(\frac{x_\tau - x}{h_x}\right) \kappa^2\left(\frac{\tau - t}{h_t}\right) E\left\{\lambda(\tau, x_\tau, \delta_\tau)\right\}\, d\tau\, dx_\tau$$

Defining $\lambda(\cdot) \equiv 0$ when $\tau > T$, we can rewrite:

$$\begin{aligned} var\left(\xi(x,t)\right) &= \tfrac{1}{h_t^2 h_x^2} \int_0^\infty \int_0^\infty \kappa^2\left(\tfrac{x_\tau - x}{h_x}\right) \kappa^2\left(\tfrac{\tau - t}{h_t}\right) E\left\{\lambda(\tau, x_\tau, \delta_\tau)\right\}\, d\tau\, dx_\tau = \\ &= \left(h_t\, h_x\right)^{-1} f(x,t) \left\{\int_0^\infty \kappa^2\left(\psi\right)\, d\psi\right\}^2 + o\left(\left(h_t\, h_x\right)^{-1}\right) \end{aligned} \tag{C.2}$$

One can note that $\xi^{(k)}(x,t)$ are independent and distributed identically as the stochastic integral generated by Poisson measure $J(\cdot)$. Therefore, for $\overline{\xi}(x,t) = \frac{1}{n}\sum\limits_{k=1}^{n} \xi^{(k)}(x,t)$ we can apply the CLT in the standard form taking into account the fact that h_t and h_x should approach 0. This suggests that:

$$\sqrt{n h_t h_x}\left(\overline{\xi}(x,t) - \xi(x,t)\right) \xrightarrow{d} N\left(0,\; f(x,t)\left\{\int_0^\infty \kappa^2\left(\psi\right)\, d\psi\right\}^2\right) \tag{C.3}$$

Asymptotic properties of the KLIC estimator

Consider a single realization of the stochastic process $\{x_\tau^{(k)}, t_\tau^{(k)}\}_{\tau\in[0,T]}$. For this realization one can write:

$$\sum_{i=1}^{I_k} \log\left[\frac{\widehat{f}(x_i^{(k)},\, t_i^{(k)})}{\widehat{f}_\theta(x_i^{(k)},\, t_i^{(k)})}\right] = \int_0^T \log\left[\frac{\widehat{f}(x_\tau^{(k)}, \tau)}{\widehat{f}_\theta(x_\tau^{(k)}, \tau)}\right]\, dJ\left(\tau,\, x_\tau^{(k)}\right),$$

where $J(\cdot)$ is a Poisson measure, generating stochastic process $\{x_\tau^{(k)}, t_\tau^{(k)}\}_{\tau\in[0,T]}$. Let ξ_1 and ξ_2 are nuisance parameters in non-parametric estimates of $f_{\theta_0}(\cdot)$ and $f_\theta(\cdot)$. Let us denote the latter stochastic integral by $\mathcal{J}^{(k)}\left(\theta,\, \xi_1,\, \xi_2\right)$.

Note now that KLIC can be written as

$$\mathcal{J}_n\left(\theta,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) = \frac{1}{n}\sum_{k=1}^{n} \mathcal{J}^{(k)}\left(\theta,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right).$$

Note that the structure of the estimator is similar to that of M-estimators, but the deterministic moment condition in this case is substituted by a stochastic moment condition. Under correct specification the estimator for θ_0 satisfies:

$$\mathcal{J}_n\left(\widehat{\theta},\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) = 0$$

with probability 1 as $n \to \infty$.

To prove the consistency and derive asymptotic properties of the obtained estimator we follow first the reasoning in [And94b]. In fact note that one can write down the mean value expansion as:

$$o_p(1) = \mathcal{J}_n\left(\widehat{\theta},\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) = \sqrt{n}\mathcal{J}_n\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) + \frac{\partial}{\partial\theta'}\mathcal{J}_n\left(\theta^*,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right)\sqrt{n}\left(\widehat{\theta} - \theta_0\right)$$

As the integral over Poisson measure is a counting integral assuming that Poisson measure does not depend on parameter θ:

$$\frac{\partial}{\partial\theta}\mathcal{J}_n\left(\theta^*,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) \xrightarrow{p} -Q = -E\left\{\int_0^T \frac{\partial f\left(x_\tau, \tau, \theta_0\right)/\partial\theta}{f\left(x_\tau, \tau, \theta_0\right)}\, dJ(\tau, x_\tau)\right\}$$

A sufficient condition for the finiteness of Q is, according to [GS79], that the integral:

$$\int_0^T \int_0^\infty \left[\frac{\partial f\left(x, \tau, \theta_0\right)/\partial\theta}{f\left(x, \tau, \theta_0\right)}\right]^2 \lambda(\tau, x)\, d\tau dx < \infty, \tag{C.4}$$

is finite.

The expression for the parameter under consideration can be written as:

$$\begin{aligned}&\sqrt{n}\left(\widehat{\theta}-\theta_0\right)=Q^{-1}\sqrt{n}\mathcal{J}_n\left(\theta_0,\,\widehat{\xi_1},\,\widehat{\xi_2}\right)+o_p(1)=\\&=Q^{-1}\left[\sqrt{n}\left(J_n\left(\theta_0,\,\widehat{\xi_1},\,\widehat{\xi_2}\right)-\mathcal{J}_n^*\left(\theta_0,\,\widehat{\xi_1},\,\widehat{\xi_2}\right)\right)+\sqrt{n}\mathcal{J}_n^*\left(\theta_0,\,\widehat{\xi_1},\,\widehat{\xi_2}\right)\right],\end{aligned}\tag{C.5}$$

where $\mathcal{J}_n^*\left(\theta,\,\xi_1,\,\xi_2\right)=\frac{1}{n}\sum\limits_{k=1}^{n}E\{\mathcal{J}^{(k)}\left(\theta,\,\xi_1,\,\xi_2\right)\}$.

As ξ_1 and ξ_2 are nuisanse parameters in the non-parametric density estimation, it has been shown in (C.3)that, similarly to standard kernel estimators as in [Sil86] such estimates are pointwise asymptotically normal. Using Fubbini theorem one can argue that $var_\xi\left[\mathcal{J}^{(k)}\left(\theta_0,\,\widehat{\xi_1},\,\widehat{\xi_2}\right)\right]=E\{\int_0^T var_\xi\left[\log\left[\frac{\widehat{f}(x_\tau,\tau)}{\widehat{f}_{\theta_0}(x_\tau,\tau)}\right]\right]\,dJ(\tau,x_\tau)\}$. The existence of this variance is justified by the existence of finite variances of $\log\left(\widehat{f}(\cdot)\right)$ and $\log\left(\widehat{f}_\theta(\cdot)\right)$. For kernel density estimates the existence of such variances is justified by the fact that $f(\cdot)$ and $f_\theta(\cdot)$ are poitwise greater than 0 on their domain. In particular, if the simulated sample is independent from the actual sample of trajectories we can write the asymptotic expression for the variance given that the density estimates are obtained from the kernel smoother with a kernel function $\kappa(\cdot)$ given assumptions A4 and A5 as:

$$\begin{aligned}&var_\xi\left[\sqrt{nh_th_x}\mathcal{J}_n^*\left(\theta_0,\,\widehat{\xi_1},\,\widehat{\xi_2}\right)\right]\\&\qquad\stackrel{n\to\infty}{\longrightarrow}2\left(\int_0^\infty\kappa^2(\psi)\,d\psi\right)^2E\left\{\int_0^T\frac{1}{f(x_\tau,\tau,\theta_0)}\,dJ(\tau,x_\tau)\right\}=\Omega_\xi\end{aligned}$$

Note that this variance is only driven by the variance in the estimation of joint density but not by the point process per se. The reason for this is that if we could have a perfect estimate of the distribution of the process x and the timing of its jumps, then KLIC under true parameter values will be equal to zero and thus the variance of the corresponding stochastic integral would be equal to zero as well. The only source of variance in $\mathcal{J}_n^*$ is therefore the error in non-parametric density estimate.

According to [GS79], a sufficient condition for the finiteness of this variance is, similarly to the above, that:

$$\int_0^T \int_0^\infty \frac{\lambda(\tau, x)}{\left(f(x, \tau, \theta_0)\right)^2}\, d\tau\, dx < \infty. \tag{C.6}$$

We can further write then that:

$$\sqrt{nh_t h_x}\mathcal{J}_n^*\left(\theta_0,\, \widehat{\xi}_1,\, \widehat{\xi}_2\right) \xrightarrow{d} N\left(0, \Omega_\xi\right).$$

Denote now $v_n(\xi) = \sqrt{n}\left(\mathcal{J}_n\left(\theta_0,\, \xi_1,\, \xi_2\right) - \mathcal{J}_n^*\left(\theta_0,\, \xi_1,\, \xi_2\right)\right)$. Under true $\xi = \xi^0$ as $\mathcal{J}^{(k)}(\cdot)$ are independent while $v_n(\xi^0) \equiv 0$.

Assuming that integral condition (C.6) holds with $\Lambda = \int_0^T \int_0^\infty \frac{\lambda(\tau,x)}{(f(x,\tau,\theta_0))^2}\, d\tau\, dx$, then:

$$E\left\{\sup_{\|\xi-\xi'\|\leq\delta} \left|\mathcal{J}^{(k)}\left(\theta_0, \xi'\right) - \mathcal{J}^{(k)}\left(\theta_0, \xi\right)\right|^2\right\} \leq 4\Lambda\delta^2,$$

by Doob's inequality. Then given mean - square convergence of non-parametric estimates of density under $\mathbf{L}^2$ norm, according to [And94b], $v_n(\cdot)$ is stochastically equicontinuous in ξ which implies that $v_n\left(\widehat{\xi}\right) - v_n\left(\xi^0\right) \xrightarrow{p} 0$.

Given these assumptions one can write the asymptotic distribution of (C.5) for the kernel density estimates:

$$\sqrt{nh_t h_x}\left(\widehat{\theta} - \theta_0\right) \xrightarrow{d} N\left(0, Q^{-1}\,\Omega_\xi\, Q^{-1}\right) \tag{C.7}$$

This establishes the fact that the obtained estimates of the vector of structural parameters θ are asymptotically normal. The finiteness of variance of such estimates is justified by the fact that as the state variable x increases, the frequency of the Poisson measure should decay faster than the density of the distribution of state variable for each moment of time.

Convergence of the MCMC estimator.

To prove convergence of the MCMC estimator consider the difference:

$$\widehat{KLIC}\,(\theta) - \widehat{KLIC}\,(\theta_0) = \tfrac{1}{n}\sum_{k=1}^{n}\int_0^T \log\left\{\tfrac{\widehat{f}(x,t,\theta)}{f(x,t,\theta_0)}\right\} dJ(t,x) = -\mathcal{J}_n\left(\theta,\,\widehat{\xi}_1,\xi_2\right). \quad \text{(C.8)}$$

The asymptotic behavior of $J(\cdot)$ has been obtained above. The only difference is that above we discussed the variance of $J^{(k)}(\theta,\,\widehat{\xi}_1,\,\widehat{\xi}_2)$ while in (C.8) we have $\mathcal{J}^{(k)}\left(\theta,\,\widehat{\xi}_1,\xi_2\right)$. By the same Fubbini theorem argument as above one can obtain $var_\xi\left[\mathcal{J}^{(k)}\left(\theta_0,\,\widehat{\xi}_1,\,\xi_2\right)\right] = E\{\int_0^T var_\xi\left[\log\left[\frac{\widehat{f}_{\theta_0}(x_\tau,\tau)}{f_{\theta_0}(x_\tau,\tau)}\right]\right]\,dJ(\tau,x_\tau)\}$.

Therefore:

$$var_\xi\left[\sqrt{nh_th_x}\mathcal{J}_n^*\left(\theta_0,\,\widehat{\xi}_1,\,\xi_2\right)\right] \overset{n\to\infty}{\longrightarrow} \left(\int_0^\infty \kappa^2(\psi)\,d\psi\right)^2 E\left\{\int_0^T \tfrac{1}{f(x_\tau,\tau,\theta_0)}\,dJ(\tau,x_\tau)\right\} = \\ = \tfrac{1}{2}\Omega_\xi$$

This in turn implies that the asymptotic behavior of MCMC estimate of θ under fixed non-parametric estimate of the data density is:

$$\sqrt{nh_th_x}\left(\widehat{\theta}_{MCMC} - \theta_0\right) \overset{d}{\longrightarrow} N\left(0,\frac{1}{2}Q^{-1}\,\Omega_\xi\,Q^{-1}\right) \quad \text{(C.9)}$$

C.6 Monte Carlo Chains of Estimated Parameters

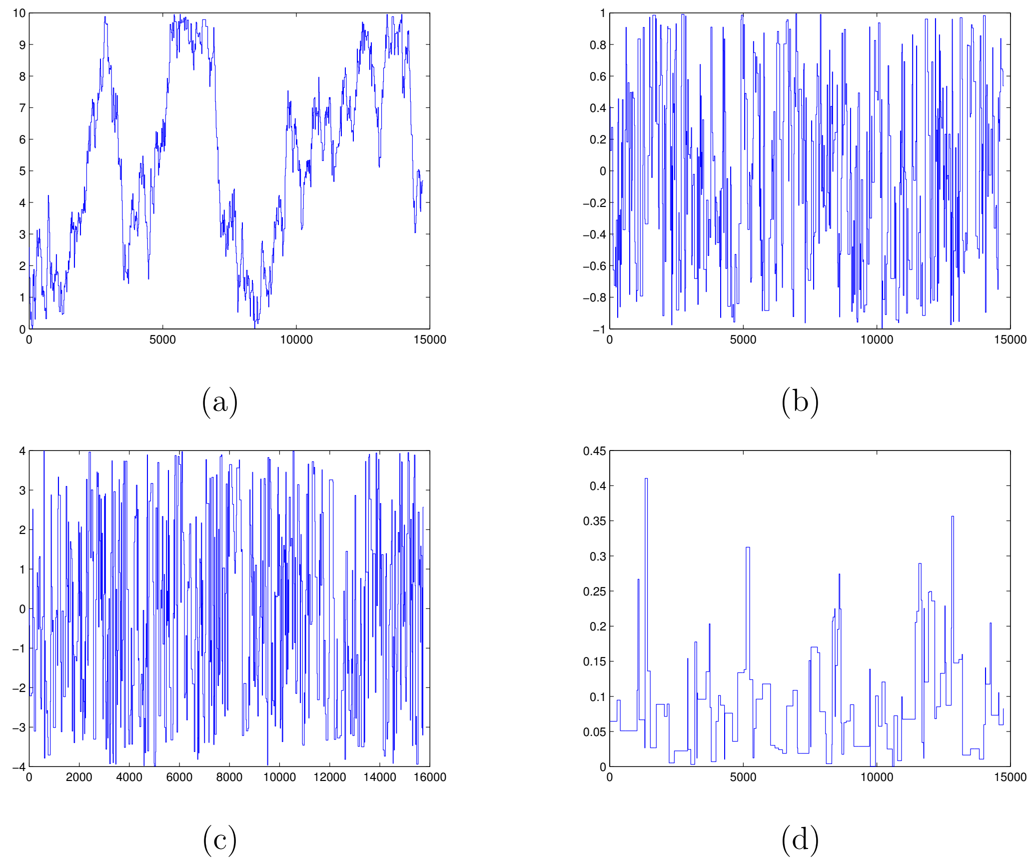


Figure C.3: Estimated parameters: a) α; b) a_2; c) a_5; d) λ_1

Bibliography

[ACHPng] Jaap H. Abbring, Pierre-Andre Chiappori, James J. Heckman, and Jean Pinquet. Adverse selection and moral hazard in insurance: Can dynamic data help to distinguish? *Journal of the European Economic Association, Papers and Proceedings*, Forthcoming.

[AH02] S. Athey and P.A. Haile. Identification of Standard Auction Models. *Econometrica*, 70(6):2107–2140, 2002.

[AHS06] D. Ackerberg, K. Hirano, and Q. Shahriar. The Buy-it-now Option, Risk Aversion and Impatience in an Empirical Model of eBay Bidding. *UCLA Working paper*, 2006.

[Ame85] T. Amemiya. *Advanced Econometrics*. Harvard University Press, 1985.

[And94a] D. Andrews. Asymptotics for semiparametric econometric models via stochastic equicontinuity. *Econometrica*, 62:43–72, 1994.

[And94b] D.W.K. Andrews. Empirical process methods in econometrics. In R. F. Engle and D.L. McFadden, editors, *Handbook of Econometrics*, volume 4. Elsevier Science, 1994.

[BB02] P. Bajari and L. Benkard. Estimation with heterogenous consumers and unobserved product characteristics: A hedonic approach. *NBER Working Paper*, 2002.

[BBL07] Patrick Bajari, Lanier Benkard, and John Levin. Estimatong dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370, 2007.

[Ber92] S. Berry. Estimation of a model of entry in the airline industry. *Econometrica*, 60(4):889–917, 1992.

[Ber94] S.T. Berry. Estimating discrete choice models of product differentiation. *RAND Journal of Economics*, pages 242–262, 1994.

[BFS03] Pierre Baldi, Paolo Frasconi, and Padhraic Smyth. *Modeling the Internet and the Web: Probabilistic Methods and Algorithms*. John Wiley & Sons, 2003.

[BH03] P. Bajari and A. Hortacsu. The winner's curse, reserve prices, and endogenous entry: empirical insights from ebay auctions. *RAND Journal of economics*, 34(2):329–355, 2003.

[BH04] P. Bajari and A. Hortacsu. Economic insights from internet auctions. *Journal of Economic Literature*, XLII:457–486, 2004.

[BHK] P. Bajari, H. Hong, and A. Khwaja. A semiparametric analysis of adverse selection and moral hazard in health insurance contracts. *Working paper*.

[BLP95] S. Berry, J. Levinsohn, and A. Pakes. Automobile price in market equilibrium. *Econometrica*, 63, 1995.

[BMSng] Michael R. Baye, John Morgan, and Patrick Scholten. Information, search and price dispersion. In T. Hendershott, editor, *Handbook on Economics and Information Systems*. Elsevier Science, Forthcoming.

[BR91] T. Bresnahan and P. Reiss. Empirical models of discrete games. *Journal of Econometrics*, 48(1):57–81, 1991.

[CDL04] D.A. Cohen, A. Dey, and T.Z. Lys. The Sarbanes Oxley Act of 2002: Implications for Compensation Structure and Risk-Taking Incentives of CEOs. *Working paper, Kellogg School of Management, Northwestern University*, 2004.

[CGPV03] S. Campo, E. Guerre, I. Perrigne, and Q. Vuong. Semiparametric Estimation of First-Price Auctions with Risk Averse Bidders. *working manuscript, University of Southern California*, 2003.

[CH04] V. Chernozhukov and H. Hong. An mcmc approach to classical estimation. *Journal of Econometrics*, 115(2):293–346, 2004.

[Che03] A. Chesher. Identification of nonseparable models. *Econometrica*, 71(5):1405–1441, 2003.

[CHS05] X. Chen, H. Hong, and M. Shum. Nonparametric likelihood ratio model selection tests between parametric and moment condition model. *Working paper*, 2005.

[Coc01] J.H. Cochrane. *Asset Pricing*. Princeton University Press, 2001.

[CSS04] Sandra Renfro Callaghan, P. Jane Saly, and Chandra Submaniam. The timing of option repricing. *The Journal of Finance*, LIX(4):1651–1676, 2004.

[CT04] F. Ciliberto and E. Tamer. Market structure and multiple equilibria in airline markets. *Working Paper*, 2004.

[DJ05] U. Doraszelski and K.L. Judd. *Avoiding the Curse of Dimensionality in Dynamic Stochastic Games*. National Bureau of Economic Research Cambridge, Mass., USA, 2005.

[DS58] N. Dunford and J. T. Schwarz. *Linear Operators. Part I: General Theory.* Wiley, 1958.

[DS03] U. Doraszelski and M. Satterthwaite. Foundations of Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity. *Manuscript, Harvard University*, 2003.

[DS05] Ulrich Doraszelski and Mark Satterthwaite. Foundations of markov-perfect indusry dynamics: Existence, purification and multiplicity. *Working paper*, 2005.

[EP95] R. Ericson and A. Pakes. Markov-Perfect Industry Dynamics: A Framework for Empirical Work. *The Review of Economic Studies*, 62(1):53–82, 1995.

[FFG69] D. Fuks, A. Fomenko, and V. Gutenmaher. *Homotopic Topology.* MSU publishing, Moscow, 1969.

[FS07] William Fuchs and Andrzej Skrzypacz. Bargaining with Arrival of New Traders. *Working paper*, 2007.

[Gal03] A.R. Gallant. Effective Calibration. *Manuscript, Fuqua School of Business, Duke University*, 2003.

[Gal04] Ronald Gallant. Efficient calibration. *Working paper*, 2004.

[GMR93] C. Gourieroux, A. Monfort, and E. Renault. Indirect inference. *Journal of Applied Econometrics*, 8:85–118, 1993.

[GPV00] E. Guerre, I. Perrigne, and Q. Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000.

[Gra72] J. Grandell. Statistical inference for doubly stochastic poisson process. In P. A. Lewis, editor, *Stochastic Point Processes: Statistical Analysis, Theory and Applications*. Wiley, 1972.

[GS79] I. I. Gihman and A. V. Skorohod. *Controlled Stochastic Processes.* Springer-Verlag, 1979.

[GT96] A.R. Gallant and G. Tauchen. Which Moments to Match? *Econometric Theory*, 12(4):657–681, 1996.

[GT02] R. Gallant and G. Tauchen. Simulated score methods and indirect inference for continuous - time models. *Working paper, Duke University*, 2002.

[GZ90] E. Gine and J. Zinn. Bootstraping general empirical measures. *The Annals of Probability*, 18(2):851–869, 1990.

[HJ99] K. Hrbacek and T. Jech. *Introduction to Set Theory*. Marcel Decker, Inc., 1999.

[HM87] Bengt Holmstrom and Paul Milgrom. Aggregation and linearity in the provision of intertemporal incentives. *Econometica*, 55(2):303–328, 1987.

[HM91] V. Hotz and R. Miller. Conditional choice probabilities and the estimation of dynamic models. *Review of Economic Studies*, 60:397–429, 1991.

[Hos04] T. Hossain. Learning by bidding. *Working paper, Hong Kong University of Science and Technology*, 2004.

[HP05] H. Hong and B. Preston. Nonnested model selection criteria. *Working paper*, 2005.

[HS83] L. P. Hansen and K. Singleton. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of Political Economy*, 92:249–265, 1983.

[HS03] H. Hong and M. Shum. Econometric models of asymmetric ascending auctions. *Journal of Econometrics*, 112:327–358, 2003.

[HS04] H. Hong and M. Shum. Rates of information aggregation in common value auctions. *Journal of Economic Theory*, 116:1–40, 2004.

[IN02] G. Imbens and W. Newey. Identification and estimation of triangular simulteneous equations models without additivity. *Working paper*, 2002.

[Izm03] S. Izmalkov. English auctions with reentry. *MIT Working paper*, 2003.

[JBP03] M. Jofre-Bonet and M. Pesendorfer. Estimation of a dynamic auction game. *Econometrica*, 71(5):1443–1489, 2003.

[JS02] W. Jank and G. Shmueli. Dynamic profiling of online auctions using curve clustering. *Working paper*, 2002.

[Kag95] J. H. Kagel. Auctions: A survey of experimental research. In *Handbook of Experimental Economics*. Princeton University Press, 1995.

[Kar86] A. F. Karr. *Point Processes and Their Statistical Inference.* Marcel Decker, Inc., 1986.

[Kos07] M.R. Kosorok. *Introduction to Empirical Processes and Semiparametric Inference.* Springer, New York, 2007.

[KP92] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer, 1992.

[Kut98] Y. Kutoyants. *Statistical Inference for Spatial Poisson Processes.* Springer-Verlag, 1998.

[Kyl85] A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1336, 1985.

[Laf92] Francine Lafontaine. Agency theory and franchising: Some empirical results. *The RAND Journal of Economics*, 23(2):263–283, 1992.

[Lag00] Ricardo Lagos. An alternative approach to search frictions. *Journal of Political Economy*, 108(5):851–873, 2000.

[LR99] D. Lucking-Reiley. Using Field Experiments to Test Equivalence between Auction Formats: Magic on the Internet. *The American Economic Review*, 89(5):1063–1080, 1999.

[LS99] Francine Lafontaine and Kathryn L. Shaw. The dynamics of franchise contracting: Evidence from panel data. *The Journal of Political Economy*, 107(5):1041–1080, 1999.

[LS01] R. Lipster and A. Shiryaev. *Statistics of Random Processes II.* Springer-Verlag, 2001.

[LSZ00] Michael L. Lemmon, James S. Schallheim, and Jaime F. Zender. Do incentives matter? managerial contracts for dual-purpose funds. *Journal of Political Economy*, 108(21):273–299, 2000.

[LV96] J.-J. Laffont and Q. Vuong. Structural analysis of auction data. *American Economic Review*, 86(2):414–420, 1996.

[MA02] M.I. Melnik and J. Alm. Does a Sellers eCommerce Reputation Matter? Evidence from eBay Auctions. *Journal of Industrial Economics*, 50(3):337–349, 2002.

[Mat03] R. Matzkin. Nonseparable estimation of nonadditive random functions. *Econometrica*, 71(5):1339–1375, 2003.

[MM00] Mary M. Margiotta and Robert Miller. Managerial compensation and the cost of moral hazard. *International Economic Review*, 41(3):669–719, 2000.

[MW04] J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall / CRC, 2004.

[Mye81] Roger Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

[Nek07a] Denis Nekipelov. Entry Deterrence and Learning Prevention on eBay. *Working paper*, 2007.

[Nek07b] Denis Nekipelov. Estimation of Continuous-Time Models of Interactions. *Working paper*, 2007.

[OR02] A. Ockenfels and A. Roth. late bidding in second price internet auctions: Theory and evidence concerning different rules for ending an auction. *Working Paper, Harvard University*, 2002.

[OY03] H. Ou-Yang. Optimal contracts in a continuous-time delegated portfolio management problem. *Review of Finacial Studies*, 16(1):173–208, 2003.

[OYGng] H. Ou-Yang and M. Guo. Incentives and performance in the presence of wealth eects and endogenous risk. *Journal of Economic Theory*, forthcoming.

[Per] I. Perrigne. Incentive regulatory contracts in public transportation: An empirical study. *Penn State University Working paper*.

[Pis90] Christopher A. Pissarides. *Equilibrium Unemployment Theory*. Oxford: Blackwell, 1990.

[PS97] M. Peters and S. Severinov. Competition among sellers who offer auctions instead of prices. *Journal of Economic Theory*, 75(1):141–179, 1997.

[PS04] M. Peters and S. Severinov. Internet auctions with many traders. *Working paper*, 2004.

[PSD03] M. Pesendorfer and P. Schmidt-Dengler. Identificication and estimation of dynamic games. *NBER Working Paper*, (9726), 2003.

[Ras03] E. Rasmusen. Getting carried away in auctions as imperfect value discovery. *Working paper, Indiana University*, 2003.

[Ras05] E. Rasmusen. Strategic implications of uncertainty over one's own private value in auctions. *Working Paper, Indiana University*, 2005.

[RC06] Christian P. Robert and George Casella. *Monte Carlo Statistical Methods*. Springer - Verlag, 2006.

[Rei06] D.H. Reiley. Field Experiments on the Effects of Reserve Prices in Auctions: More Magic on the Internet. *RAND Journal of Economics*, 37(1):195–211, 2006.

[Rib02] L.E. Ribstein. Market vs. Regulatory Responses to Corporate Fraud: A Critique of the Sarbanes-Oxley Act of 2002. *Journal of Corporation Law*, 28(1):1–67, 2002.

[Ric96] J. Richardson. Vertical Integration and Rapid Response in Fashion Apparel. *Organization Science*, 7(4):400–412, 1996.

[Rom05] R. Romano. The Sarbanes-Oxley Act and the Making of Quack Corporate Governance. *Yale Law Journal*, 114(7):1521–1613, 2005.

[Rus87] J. Rust. Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica*, 55(5):999–1033, 1987.

[Rya05] S. Ryan. The Costs of Environmental Regulation in a Concentrated Industry. *MIT Working paper*, 2005.

[San07] Y. Sannikov. Games with Imperfectly Observable Actions in Continuous Time. *Econometrica*, 75(5):1285–1329, 2007.

[Ser80] R.J. Serfling. *Approximation theorems of mathematical statistics*. Wiley New York, 1980.

[Shi99] A.N. Shiryaev. *Essentials of stochastic finance*. World Scientific River Edge, NJ, 1999.

[Sil86] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986.

[SM91] D. L. Snyder and M. I. Miller. *Random Point Processes in Time and Space*. Springer-Verlag, 2 edition, 1991.

[Spr81] G.B. Sproles. Analyzing Fashion Life Cycles: Principles and Perspectives. *Journal of Marketing*, 45(4):116–124, 1981.

[Spr85] G.B. Sproles. Behavioral science theories of fashion. In M.R. Solomon, editor, *The Psychology of Fashion*. Lexington Books, 1985.

[SRJ03] G. Shmueli, R. Russo, and W. Jank. Modeling bid arrivals in online auctions. *Technical report*, 2003.

[SS93] H. Shäller and J. Sung. The first order approach to the continuous-time principal-agent problem with exponential utility. *Journal of Economic Theory*, 61:331–371, 1993.

[Sun95] J. Sung. Linearity with project selection and controlable diffusion rate in continuous-time principal-agent problems. *RAND Journal of Economics*, 26:720–743, 1995.

[VS77] J. H. Van Schuppen. Filtering, prediction and smoothing for counting process observations, a martingale approach. *SIAM Journal of Mathematics*, 32(3):552–570, 1977.

[Vuo89] Q. Vuong. Likelihood - ratio tests for model selection and non - nested hypotheses. *Econometrica*, 57:307–333, 1989.

[WBVR05] G. Weintraub, C.L. Benkard, and B. Van Roy. *Markov Perfect Industry Dynamics with Many Firms*. National Bureau of Economic Research Cambridge, Mass., USA, 2005.

[Wil04] Noah Williams. On dynamic principal-agent problems in continuous time. *Princeton University Working paper*, 2004.

[WJS04] S. Wang, W. Jank, and G. Shmueli. Forecasting ebay's online auction prices using functional data analysis. *Working paper*, 2004.

[Yin03] Pai-Ling Yin. Information dispersion and auction prices. *Working paper, Stanford University*, 2003.

[Zha05] I.X. Zhang. Economic consequences of the Sarbanes-Oxley Act of 2002. *Working paper, William E. Simon Graduate School of Business Administration, University of Rochester*, 2005.

Biography

Denis Nekipelov was born in Russia, on October 19, 1980. He earned his M.S. in applied physics and mathematics with distinctions in June 2003, and M.A. in economics Cum Laude from the New Economic School, Moscow, in July 2003. Denis began his graduate studies at Duke University in August 2003. He was the recipient of H. Gregg Lewis fellowship from 2003 to 2004, Graduate school summer research fellowship in 2007, Micro-Incentives Research Center (SSRI) mini-grant in 2006, Graduate school Conference Travel Award in 2005, 2006, and 2007, and Department of Economics summer research fellowship in 2006.