Essays in Industrial Organization:
Umbrella Branding, Non-Binding Auctions and Opaqueness of the Patent System

Inaugural-Dissertation

zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

an der Ludwig-Maximilians-Universität München

2013

vorgelegt von

Sebastian Stoll

Referent:	Prof. Dr. Dietmar Harhoff
Korreferent:	Prof. Dr. Monika Schnitzer
Promotionsabschlussberatung:	14. Mai 2014

Datum der mündlichen Prüfung: 07. Mai 2014

Namen der Berichterstatter: Dietmar Harhoff, Monika Schnitzer, Gregor Zöttl

To my family.

Acknowledgements

First and foremost, I would like to thank my supervisor Dietmar Harhoff for his inspiration, his advice, and for the collaboration on chapter three. Also, despite his many other duties and obligations, he always offered invaluable support on matters reaching far beyond the scope of writing this dissertation. I am truly grateful to him for this. Second, I am very much indebted to Gregor Zöttl, with whom I collaborated on chapter two of this dissertation. In many discussions, Gregor taught me how to write an article in the field of economics and sharpened my economic intuition. Above all, I had a good time working closely with him. In addition, he agreed to join my thesis committee as third supervisor. I would also like to thank Monika Schnitzer. Besides kindly agreeing to join my thesis committee as second supervisor, she offered me the possibility to present my work in her internal seminar and provided me with helpful comments and suggestions. Furthermore, I gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) through GRK 801.

I was very fortunate to stay at the London School of Economics for one semester as visiting research student. I am grateful to the Department of Economics for its support and hospitality. From the inspiring research environment there I drew valuable inspirations for my work. I would like to thank Dietmar Harhoff and Monika Schnitzer for their support in arranging my stay at the London School of Economics.

More than anybody I would like to thank my family and my friends, both my old ones and the ones I made during my time in Munich. Without your support I would not be where I am now. And, most importantly: I know that I can always rely on you all to remind me of what in the end is really important in life!

Sebastian Stoll
Munich, December 2013

Contents

Preface

"I finish by expressing my hope that Industrial Organization returns to be a much more empirically-oriented discipline, guided, of course, by sound theory. We all have much to do to ensure that antitrust repeats the successes and avoids the mistakes of its past."

Timothy J. Muris, 2003.

The above statement finishes a speech Timothy J. Muris gave in January 2003 as Chairman of the U.S. Federal Trade Commission in front of an auditorium of antitrust professionals.[1] It reflects the development competition policy took in the last twenty years: From a form-based assessment of competitive behavior, which deems certain business practices to be anticompetitive per se, to an effects-based assessment, which evaluates the positive and negative competitive effects of a given business practice in a case-specific manner and weighs these effects against each other. This change in the focus of competition policy expresses itself by a rapidly increasing influence of economists in competition authorities - in the EU Directorate-General for Competition, from the early 1990s to 2008 the ratio of economists to lawyers rose from one to seven to one to two.[2] The increasing importance of economics in competition cases, and this is at the core of the quote I put at the beginning of this preface, in turn makes an empirically well-founded understanding of the inner workings of different markets all the more important.

This thesis sets out to expand the understanding of different markets by analyzing three institutions: In the first chapter, I demonstrate that umbrella brands, which mark several products as being from one product family, induce state dependence

[1]The speech is titled "Improving the Economic Foundations of Competition Policy" and was given at George Mason University Law Review's Winter Antitrust Symposium.

[2]Lowri Evans, the then Deputy Director General of the EU DG Competition, pointed this out in his 2008 speech "The role of economics in modern competition policy". It was given at the International League of Competition Law Congress 2008 in Hamburg.

in consumers' purchasing behavior over time. The second chapter, which is joint work with Gregor Zöttl, shows that the benefit of information dissemination in non-binding procurement auctions depends critically on the weight buyers attach to bidders' non-price characteristics. The third chapter, which is joint work with Dietmar Harhoff, exploits the concealment of a signal of patent value by the European Patent Office in 2001 to give evidence that the European patent system is inherently opaque. As the quote at the beginning of the preface asks for, the analysis in all three chapters is guided by theory but in the end all main findings are based on empirical data. In the long run the insights presented in this thesis will hopefully contribute to a better assessment of markets from a competition policy perspective. In the following, each chapter is outlined in turn. The order of the chapters corresponds to the order of their inception.

The first chapter of this thesis deals with a strategic instrument which - despite being commonly used by firms - has not received a lot of attention in the economic literature yet. This instrument is umbrella branding. Umbrella branding means the use of recurring brand elements on different products so that consumers perceive these products as being from one product family. I focus my analysis of umbrella branding on its implications on consumers' over-time purchasing behavior. In particular, I ask whether umbrella branding induces state dependence with respect to product families. That is: Are consumers who switch from a product they previously purchased to another one more likely to switch to a product under the same umbrella brand? I find that marking several products as related by the use of an umbrella brand indeed causes consumers to stick to this family of products. The important point to stress here is that this sticky behavior is not simply rooted in consumers' preferences but actually due to the fact that umbrella branding induces switching costs. Also, loyalty to a product family induced by umbrella branding is of an economically significant size - on average, it amounts to around 25% of the mean product price.

Methodologically, at the core of identification of state dependence in consumers' purchasing behavior with respect to umbrella brands lies the application of a discrete choice model to a panel of households from which data about their grocery purchases was collected. The critical point in identification is the separation of preference heterogeneity from structural state dependence. In other words, if the discrete choice model did not properly account for the fact that some consumers might prefer one umbrella brand over the others, then it would falsely attribute repurchases of this umbrella brand by these consumers to structural state dependence.

Thus, the choice model has to be able to flexibly capture different forms of preference heterogeneity among consumers. In specifying a flexible model, I follow Dube et al. (2010), who proposed to represent preference heterogeneity by a mixture of normals distribution. The mixture of normals distribution is scalable to capture any kind of preference distribution among consumers while still being computationally feasible. With controls for preference heterogeneity in place, my model identifies state dependence in consumers' umbrella brand choices. The vast amount of data available allows me to explore this result further. I find evidence that the observed umbrella brand loyalty is rooted in psychological switching costs and not search or learning cost, and that it is present both within and across product categories.

These findings add to the economic understanding of the practice of umbrella branding. So far, economic articles on umbrella branding, like Cabral (2000, 2009) and Hakenes and Peitz (2008), focus on the function of umbrella brands as quality signals: Roughly put, the basic assumption there is that consumers expect the qualities of products which are assembled under an umbrella brand to be correlated. Then, in equilibrium firms choose high qualities for all products under an umbrella brand, and consumers accordingly expect products under an umbrella brand to be of high quality. Given the assumption that product quality is kind of a fixed cost investment and does not (or only slowly) change over time, the dynamic implications of this view are restricted to an initial learning period. As soon as consumers are experienced with regard to umbrella brands (that is, the level of product quality these stand for), umbrella branding should no longer have an influence on consumers' purchasing decisions. However, due to the length of the consumer panel I have available I can show that also experienced consumers exhibit inertia in their umbrella brand choices. This finding gives evidence to the fact that state dependence in umbrella brand choice is neither rooted in learning nor search costs. Thus, in order to fully assess the role the practice of umbrella branding plays in competition, the view of umbrella brands as quality signals has to be complemented by the view of umbrella brands as product characteristics which induce structural switching costs.

The second chapter is joint work with Gregor Zöttl. In this chapter we analyze open non-binding auctions, an auction format which dominates the rapidly expanding online procurement market. Basically, open non-binding auctions are of a very simple structure: Initially, a buyer publishes a description of the product or job he wants to procure. During a predefined period of time, bidders can then put forward price quotes. These price quotes are publicly visible and can be changed anytime. At the end of the bidding period, the buyer freely decides for one of the participating bidders. Despite their very simple structure and their increasing importance, open non-binding auctions are not yet well understood. The reason is that open non-binding auctions do not fit into the traditional and well-established framework of auction theory and thus elude the treatment with standard tools.

We develop a theoretical framework to describe open non-binding auctions. We do this to assess the effects of availability of non-price information to the bidders - as bidding takes place in an online environment, the distribution of information is easy to manipulate by the creator of an auction, and we indeed observe different informational arrangements in the field. In particular, we compare an informational arrangement where bidders are informed about their rivals' non-price characteristics to one where this information is concealed from them. The questions we ask are: Under what conditions does a buyer prefer to conceal information about their rivals' non-price characteristics from the bidders? And: How large can welfare effects of a change in the information structure expected to be in the field?

From our theoretical framework, we find that whether the buyer prefers to disclose or to conceal non-price information depends on how the buyer weighs bidders' non-price characteristics against bidders' prices. In case the buyer puts a lot of weight on bidders' non-price characteristics, he is better off when he conceals non-price information. On the other hand, in case he puts only small weight on bidders' non-price characteristics, the buyer is better off when he discloses non-price information. The intuition is simple: If bidders are mainly differentiated by their non-price characteristics, then concealment of this fact makes them appear more similar to each other, which intensifies competition and thus leads to lower price quotes. However, if bidders mainly compete on prices, then disclosure of information about their non-price characteristics softens the advantage of bidders who are able to offer low prices due to low costs, which in turn again intensifies competition and decreases overall prices.[3]

We use data on open non-binding auctions from a large European procurement platform to quantitatively assess the effects of a change in information structure. The information structure on this platform is such that bidders' non-price characteristics are public information. First, we establish that bidding behavior is in line with our theoretical predictions for the case bidders are informed about their rivals' non-price characteristics. We then employ our theoretical framework to perform a counterfactual analysis. The first step in our counterfactual analysis is to derive

[3]The assumption here is that lower costs correlate with less favorable non-price characteristics.

estimates of bidders' costs using our framework for the case bidders are informed about each others' non-price characteristics. We use these cost estimates as inputs for our framework for the case non-price information is concealed to derive counterfactual outcomes. According to our intuition, for auction categories where bidders' skills are relatively important to the buyer we expect buyers' welfare to increase by up to ten percent in case non-price is concealed, while for categories where bidders' skills are of little importance we expect a decrease of up to ten percent.

As indicated above, despite the rapidly increasing importance of the open non-binding auction format for both firms' and private persons' procurement activities, the economic literature on them is still very scarce. To our knowledge, so far there is only one article which explicitly deals with the structuring of information in open non-binding auctions, which is Haruvy and Katok (2013). In this article the authors report the results of an experimental study. For their setup they find that buyer surplus increases when information about their rivals' non-price characteristics is concealed from the bidders. Our contribution places their result into a broader context by demonstrating that the effect of a change in information structure is not unambiguous but depends on the precise characteristics of the auction under consideration. Thus, any recommendation on the information structure of open non-binding auctions has to be based upon a thorough empirical analysis of the specific auction environment.

The third chapter of this dissertation, which is joint work with Dietmar Harhoff, sheds new light on the fundamental tradeoff of the patent system - that is, the granting of exclusion rights in exchange to disclosure of technical knowledge. Many authors in the patent literature take it as given that the patent system fully discloses technical knowledge and concentrate on the incentive structures arising from the granting of exclusion rights. We, however, differ from the widespread belief that the patent system fully discloses critical information about the innovations protected by patents. Instead, we argue that the patent system is indeed highly opaque with respect to the technical and economical value of a patent. We base our position on data from a quasi-experimental setting: In December 2001, the European Patent Office (EPO) changed its information policy regarding requests for accelerated patent examinations. While before December 2001 information about whether a patent applicant requested accelerated examination was publicly available, this information was treated as confidential afterwards. In reaction to the 2001 concealment of acceleration information, in our data we observe the behavior of patent applicants and their rivals to change in a way consistent with our assumption that the patent system is opaque with respect to patent value. That is, it seems that the information conventionally generated by the EPO is not sufficient to allow rivals to identify a patent's actual contribution.

In particular, we start out by developing a theoretical model of the patent application and opposition process. We model this process as a dynamic two-stage game, where first a patent applicant draws either a high- or low-value patent and then decides whether to request accelerated patent examination. Second, a rival of the applicant decides whether to oppose the patent. We use this basic structure to compare outcomes in case the patent system is fully transparent with respect to patent value to outcomes in case it is opaque. In case the patent system is opaque, we in addition compare outcomes in case the applicant's acceleration decision is disclosed to outcomes in case it is concealed. From this framework we derive predictions about the way the behavior of applicants and rivals should change in reaction to the EPO's 2001 decision to conceal information about acceleration requests: Whereas in case of a transparent patent system we expect to observe no changes in behavior, in case of an opaque system we expect the rate of acceleration requests to increase and the rate of oppositions to decrease. In our data, we indeed see the latter predictions confirmed: The frequency of acceleration requests is significantly higher after the EPO's 2001 policy change than it is before, and the frequency of oppositions is significantly lower. That is, the data supports our presumption that the European patent system is opaque with respect to the value of patents. Our main finding therefore is that the conventional data generated by the EPO is not suited to identify competing approaches and firms easily, and that thus the patent system is probably limited as a source of information.

Our model allows us to take a first step towards a welfare assessment of the result that the European patent system is opaque with respect to patent value. Maybe surprisingly, we find that opaqueness with respect to patent value might be beneficial for the aggregate welfare of applicants and rivals. However, this is only a partial welfare result, as our model focuses on the parties directly involved in the application and opposition process and is agnostic with respect to the implications of opaqueness of the patent system for third parties and thus for the progress of innovation in society. With respect to an assessment of the EPO's 2001 policy change, our model shows that the welfare implications of concealment of the acceleration signal critically hinge on how strongly the value of a patent increases in case its examination is accelerated. Unfortunately, the issue of patent acceleration seems not to have received significant attention in the patent literature so far. Thus, without further research into the issue of patent acceleration well-founded statements about welfare changes due to the 2001 concealment of the acceleration signal cannot be made.

The three chapters of this thesis are self-contained and include their own introductions and appendices. Hence, each chapter can be read on its own. References for all three chapters are listed at the end of this thesis.

Chapter I

Umbrella Branding and Consumer Inertia

I.1 Introduction

Umbrella brands mark products as being from the same product family by the use of recurring brand elements. As an example: The umbrella brand “Colgate” establishes a link between a “Colgate Active Angle” toothbrush and a “Colgate Total” toothbrush. In this chapter I empirically analyze how umbrella brands influence households in their over-time purchasing behavior. In particular, I am interested in whether there is inertia in households’ umbrella brand choices. Inertia in umbrella brand choice means that, conditional on a product change, ceteris paribus a household is more likely to switch to a product under the previously purchased umbrella brand than to a product under another umbrella brand.

I show that there is significant inertia in households’ umbrella brand choices. By making use of the length of the household panel I have available, I demonstrate that inertia in umbrella brand choice can be rationalized by the existence of structural (respectively psychological) switching costs, but not by search or learning costs. In addition, availability of data from different product categories allows me to show that inertia in umbrella brand choice exists both within and across product categories. These results add to the economic literature on umbrella branding, which, despite of the prevalence of umbrella branding in everyday life, is still quite scarce: So far, umbrella brands have primarily been understood as quality signals. The finding that firms can use umbrella brands to induce structural switching costs sheds a new light on the practice of umbrella branding, especially with respect to the assessment of competitive behavior.

Specifically, I develop a dynamic discrete choice model to determine whether there is inertia in households' umbrella brand choices. In this model, besides product prices and product-specific intercepts, I include covariates controlling for households' purchase histories: In particular, I include a covariate which controls for a households' previous umbrella brand choice. The coefficient on this covariate captures inertia in households' umbrella brand choices. That is: Given that a household purchased a product under a certain umbrella brand at his previous shopping trip, the coefficient measures whether (respectively how much) the utilities of products which are not under this umbrella brand are lowered at the household's present shopping trip. It is identified by changes in households' product choices which are induced by price variations.

A common concern in the empirical literature on state dependence in households' choices is that a model spuriously identifies inertia in households' choice behavior. Reasons for spurious identification of choice inertia are unobserved household heterogeneity or unobserved correlations in households' tastes for the products considered (compare for example Dube et al., 2010). To address this concern, I specify the model coefficients to follow mixtures of multivariate normal distributions, and I allow the product specific intercepts to be correlated. That way I flexibly control both for various forms of unobserved household heterogeneity and for unobserved correlations in a household's product tastes.

I estimate my model on household panel data from the IRI marketing dataset. Specifically, I concentrate on data on toothbrush purchases. The toothbrush category is well-suited for my research purposes: In this category, the use of umbrella brands is common, and households' shopping behavior fits the general assumptions of a discrete choice framework very well. Also, availability of data on purchases in the toothpaste category allows me to examine cross-category effects of umbrella branding. The data on consumers' shopping trips stems from two large metropolitan areas in the US, Eau Claire, Wisconsin, and Pittsfield, Massachusetts, and comprises the years 2001 to 2005.

I use a Bayesian estimation approach to derive posterior distributions of my model coefficients. For computation of the posterior distributions I employ a Markov Chain Monte Carlo algorithm. I find that if the average consumer changes from the product he purchased in the previous period to some other product, then he is indifferent between a product being under the same umbrella as the previously purchased one and a product being under another umbrella but with a by \$0.69 lower price. (The simplifying assumption here is that the products the household can choose from are physically identical.) Put in slightly other terms, the dollar value of umbrella brand loyalty on average amounts to 25% of the mean product price.

One concern with these results is that the lagged umbrella brand choice coefficient does not capture inertia in households' umbrella brand choices but in fact simply picks up correlations in households' product tastes. To counter this concern I estimate my model with and without the lagged umbrella brand choice coefficient. Inclusion of the lagged umbrella brand choice coefficient significantly increases the posterior probability of my model, which evidences that the lagged umbrella brand choice coefficient actually captures dynamics in households' choice processes. Furthermore, my estimation results are robust under different prior specifications, and do not change when I account for the possibility that besides by prices households' decisions might be influenced by marketing activities on and above store-level. In summary, there is strong evidence that my finding of inertia in umbrella brand choice is not spurious but captures actual regularities in the dynamic choice behavior of households.

The length of my observation period allows me to explore the behavioral underpinnings of my finding of structural inertia in households' umbrella brand choices. In particular, I split my observation period in two subperiods and estimate my choice model only for the late subperiod and only on a subsample of experienced households. Experienced households are those which in the early subperiod of my observation period visited all stores in my sample and purchased products under all umbrella brands. For the subsample of experienced households I still observe inertia in umbrella brand choice. This excludes search and learning costs as possible explanations for consumer inertia in umbrella brand choice. Thus, it seems that inertia in umbrella brand choice is caused by the existence of structural (or psychological) switching costs.

In practice, an umbrella brand often not only assembles products from one but from several product categories. In my dataset, besides on toothbrush purchases I have available data on toothpaste purchases, and there are umbrella brands which are present in both the toothbrush and the toothpaste category. This allows me to analyze whether besides within a product category inertia in umbrella brand choice also exists across product categories. I find that households' decisions for an umbrella brand in the toothbrush category are significantly affected by whether they previously purchased a toothpaste product under this umbrella brand. Hence, inertia in umbrella brand choice is a phenomenon which exists both within and across product categories.

My work basically adds to two strands of literature: One which empirically deals with state dependence in households' purchasing behavior, and another which tries to understand the economic rationale behind the use of umbrella brands. The literature on state dependence in households' purchasing behavior goes back to Frank (1962) and Massy (1966). From then on, inertia in product choice has been well documented by several authors, among those for example Keane (1997), Erdem (1998) and Seetharaman et al. (1999). In a recent article, Dube et al. (2010) use a discrete choice model with a flexible heterogeneity specification to show that inertia in product choice is indeed a structural phenomenon and not caused by unobserved heterogeneity in consumer preferences. They additionally explore possible economic explanations for the existence of inertia in product choice and find that inertia in product choice is most likely caused by structural (or psychological) switching costs and not search or learning costs. This chapter is closest to their article.

In contrast to Dube et al. (2010), however, I do not treat products as single, disconnected entities. Instead, I explicitly take into account that often products are marked as being related by an umbrella brand. To my knowledge, there is only one other article on household inertia which also explicitly accounts for the fact that products might be connected by an umbrella brand: For the yogurt product category, Pavlidis and Ellickson (2012) show that households exhibit inertia in umbrella brand choice. However, while their article is mainly concerned with the strategic pricing issues arising from the existence of state dependence with respect to umbrella brands, my contribution focuses on the analysis of whether the observed inertia in households' umbrella brand choices is indeed structural and not spurious, the isolation of the effect of inertia in umbrella brand choice from that of inertia in product or brand choice, and the exploration of the mechanisms behind the phenomenon of household inertia in umbrella brand choice. In summary, I contribute to the literature on state dependence in households' choice behavior by demonstrating that choice inertia is not restricted to single products but also present with respect to product families, and by analyzing the mechanics behind the observation of choice inertia with respect to product families.

The other strand of literature this chapter (more indirectly) adds to is that of the economic theory of umbrella branding: The theoretical literature on umbrella branding offers different economic explanations for its existence. One, put forward for example by Andersson (2002), is economies of scope: several products under one umbrella brand can easily be promoted by a single marketing campaign. Another, which seems to be in the focus of recent research on umbrella branding, is the possibility to signal product quality by the use of umbrella brands. Papers which theoretically analyze the role of umbrella brands as signals for product quality are for example Wernerfelt (1988), Choi (1998), Cabral (2000, 2009) and Hakenes and Peitz (2008). There is some supporting empirical evidence indicating that households' preferences for products linked by an umbrella brand are correlated (Sullivan, 1990; Erdem, 1998; Seetharaman et al., 1999; Erdem and Sun, 2002; Balachander and Ghose, 2003).

In contrast to these contributions, however, I show that even when controlling for preference correlations a consumer's previous-period umbrella brand choice has a direct influence on his present shopping decision. That is, I contribute to this strand of literature by offering a third rationale for the existence of umbrella brands: As conditional on a product change households incur additional switching costs if they also change umbrella brand, firms can use umbrella brands as lock-in devices for families of products which are possibly located in several product categories. Other than the rationales for umbrella branding established in the economics literature so far, this new rationale has implications directly related to market structure and competition between firms. It explains for example the observation of product proliferation under established umbrella brands: In analogy to Klemperer (1995), households who value variety and who incur umbrella brand-specific switching costs prefer umbrella brands which assemble a lot of different products over umbrella brands which assemble only a few.

The next section introduces the discrete choice model I use to examine whether there is inertia in households' umbrella brand choice. Section I.3 describes the household panel data I apply this model on. Section I.4 presents my main estimation result: There seems to be significant inertia in households' umbrella brand choice. In section I.5 I show that this result is not spurious due to misspecified household preferences and that it is robust under different model and prior specifications. Section I.6 explores the behavioral underpinnings of the observation that households have a tendency to stay with an umbrella brand. In section I.7 I show that inertia in umbrella brand choice occurs both within and across product categories. Section I.8 concludes.

I.2 The Choice Model and its Econometric Specification

■ **Model.** In the following I will use the term *product* to denote the entity a household can actually take away from the shelfs of a grocery store. A *brand* is a symbol applied to a physical product to distinct this physical product from other (possibly physically identical) products. An *umbrella brand* is a symbol common to several brands which marks these brands as being from the same brand family.[1] I assume that there are J products, which are assembled under B brands, which in turn are assembled under U umbrella brands.

I assume that there are H households h, each observed at T_h storeweeks t. The term storeweek denotes the visit of a certain store in a certain week.[2] I assume the storeweeks to be exogenously given. At a storeweek t a household h has the choice among J products.[3] The utilites household h derives from each of the products at shopping occasion t are

$$\begin{aligned} u_{0t}^h &= \epsilon_{0t}^h, \\ u_{1t}^h &= \alpha_1^h + \eta^h p_{1t} + \beta^h \mathbb{1}_{B(1)\neq B_{t-1}} + \gamma^h \mathbb{1}_{U(1)\neq U_{t-1}} + \epsilon_{1t}^h, \\ &\vdots \\ u_{Jt}^h &= \alpha_J^h + \eta^h p_{Jt} + \beta^h \mathbb{1}_{B(J)\neq B_{t-1}} + \gamma^h \mathbb{1}_{U(J)\neq U_{t-1}} + \epsilon_{Jt}^h. \end{aligned} \qquad \text{(I.1)}$$

The index 0 denotes the household's outside option, that is the decision of the household not to purchase any of the J products. The α_j^h are product specific intercepts

[1] As an example, a Colgate Active Angle toothbrush with a full head and a soft bristle is a product, "Colgate Active Angle" is its brand, and "Colgate" its umbrella brand. Note that in the literature the use of the terms product, brand and umbrella brand is not consistent. For example, Dube et al. (2010) use the term brand where I use the term product, while Erdem (1998) uses the term brand where I use the term umbrella brand.

[2] A storeweek is similar but not synonymous to a shopping trip, as at a storeweek t a household could have visited the respective store several times the respective week. I use the term storeweek because the finest temporal resolution in our data amounts to one week.

[3] Note that the extension of the model to the case that at storeweek t household h can only choose among a subset of the J products is straightforward. In fact, to derive my estimation results it is crucial to use a model which accounts for varying choice sets. However, as the extension of the model to the case of a varying choice sets just brings about more involved notation, for the sake of exposition I present the model for the case of a constant choice set.

which capture household h's intrinsic preferences for product j. p_{jt} is the price of product j at storeweek t. The price coefficient η^h measures the change in the utility household h derives from product j if its price is altered by one dollar. $\mathbb{1}_{B(j)\neq B_{t-1}}$ is an indicator variable which establishes a connection between a household's current and its previous brand choice. It equals one if the brand of product j is equal to the brand of the product the household purchased in the previous storeweek and zero otherwise. Analogous, $\mathbb{1}_{U(j)\neq U_{t-1}}$ is an indicator variable which establishes a connection between a household's current and its previous umbrella brand choice. It equals one if the umbrella of product j is equal to the umbrella of the product the household purchased in the previous storeweek and zero otherwise.[4] Accordingly, the coefficients β^h and γ^h measure how the utility household h receives from product j is altered if product j is under a different brand respectively a different umbrella brand than the product household h chose in the previous period. If β^h is negative, household h exhibits inertia in brand choice. If γ^h is negative, household h exhibits inertia in umbrella brand choice. The ϵ^h_{jt} are error terms capturing (storeweek-dependent) deviations in household h's behavior. Note that I allow the model coefficients $\boldsymbol{\theta}_h \equiv (\alpha^h_1, ..., \alpha^h_J, \eta^h, \beta^h, \gamma^h) \equiv (\boldsymbol{\alpha}^h, \eta^h, \beta^h, \gamma^h)$ to be household-specific.

I assume that at storeweek t a household h makes a discrete choice among the J products. That is, at storeweek t household h is assumed to choose exactly one of the J products or the outside option. The household will choose the product which maximizes its expected utility at storeweek t. Thus, household h's choice problem at storeweek t is

$$\underset{j\in\{0,1,...,J\}}{\operatorname{argmax}}\ u^h_{jt}. \tag{I.2}$$

■ **Identification.** The coefficients β^h and γ^h in model (I.1) account for inertia in households' choice behavior: If there is inertia in brand choice, the coefficient β^h will be smaller than zero. If conditional on a brand change there is inertia in umbrella brand choice, the coefficient γ^h will be smaller than zero.

Key for identification of the coefficients β^h and γ^h is variation in product prices and households' choices. The following simple example illustrates the basic identification mechanism: Assume that at storeweek one a given household chooses product one.

[4]Note that I also could include a covariate accounting for a household's previous-period product choice. However, as this covariate turned out to be of no influence on households' purchasing decisions, for the sake of clarity I did not include it in the present model.

At storeweek two all prices stay the same, except that of product two, which is decreased so far that the household now chooses product two. Product two shall be under both another brand and another umbrella brand than product one. At storeweek three the price of product two is increased again to a level slightly above its price in storeweek one. Again, the prices of all other products shall stay the same.

Now there are three different possible choices of the household in storeweek three, each of which having different implications regarding the existence of inertia in households' choice behavior: First, the household might choose product one again. That would either mean that there is no inertia or that it is too small to affect the household's behavior. Second, the household might stay with product two. That would mean that there is inertia with respect to brands. Third, the household might choose neither product one nor product two but might change to a third product which is under the same umbrella as product two. That would indicate that there is inertia with respect to umbrella brands.[5]

For the sake of exposition, in the example above I implicitly assumed that there are no idiosyncrasies in the household's decisions (that is $\epsilon_{jt}^h = 0$ for all t and j). This assumption is of no effect to the general logic of the example. If I allowed for non-zero error terms, and if I assumed that the error terms were independently and identically distributed and uncorrelated with product prices, then statements about the existence of switching costs could be deduced from the "average behavior" of the household. That is, if the household were repeatedly exposed to the pricing dynamics above, and if in the third period of each pricing sequence the household in "almost all" cases chose to switch to a product under the same umbrella as product two, then that would hint to the existence of choice inertia with respect to umbrella brands.

Like there are pricing patterns which are informative about the existence of choice inertia, there are also pricing patterns which are informative about the households' preferences: As a simple example, if at storeweek t a household changes from some other product to product one, we know for sure that $\alpha_1^h - \eta^h p_{1t} \geq \alpha_j^h - \eta^h p_{jt}$ for all products j which are under the same umbrella as product one. As product

[5]Note that by varying the price of product two in storeweek three also the size of the inertia coefficient relative to that of the price coefficient can be identified.

prices are observable, we thus have information about the relationships between the household's preferences for all products under the same umbrella as product one.[6]

Put together, the household-specific coefficients in model (I.1) are identified if three conditions hold: First, unobserved influences on the utilities households derive from product purchases must not be correlated with each other (both in the product- and the time-dimension) or with product prices. Second, there has to be sufficient price variation in the data. Third, there has to be sufficient variation in households' choices.

■ **Econometric specification.** The above discussion revealed that it is possible to draw conclusions about the coefficients of model (I.1) from the observation of households' choice behavior when households are faced with varying product prices. As the amount of data available is not sufficient to make inferences about the coefficients of model (I.1) without further assumptions, I have to make parametric assumptions.

Household panels are usually short, which means that meaningful coefficient estimates on the level of a single household cannot be derived. Instead, I retreat to assumptions about the distribution of the model coefficients in the population of households. As already shown for example by Dube et al. (2010), it is crucial for the identification of choice inertia that the specified distribution of the model coefficients captures heterogeneity in the preferences of the households sufficiently well.

To make this clear, assume that there is no inertia whatsoever and that there are two equally large groups of households in my data: One which likes products from umbrella one but dislikes products from umbrella two, and another which dislikes products from umbrella one but likes products from umbrella two. As a result, simply because of these differences in preferences group one will on average stay with products from umbrella one, whereas group two will on average stay with products from umbrella two. Now, if I specified the distribution of the household coefficients to be degenerate (meaning that each household has the same preferences for all products), the only possible explanation for this behavior would be the existence

[6] As for the example above, which illustrates the identification of switching costs, the implicit assumption made here is that $\epsilon_{jt}^{h} = 0$ for all t and j. The generalization to non-zero error terms again is straightforward: If I allowed for non-zero error terms, and if I assumed that the error terms were independently and identically distributed and uncorrelated with product prices, then statements about a household's preferences could be deduced from the "average behavior" of the household.

of choice inertia with respect to umbrellas. That is, econometric specifications of model (I.1) which do not account for heterogeneity in households' preferences might lead to spurious detection of choice inertia.

To avoid spurious detection of choice inertia I follow Dube et al. (2010) and specify my model coefficients $\boldsymbol{\theta}_h = (\boldsymbol{\alpha}^h, \eta^h, \beta^h, \gamma^h)$ to follow a mixture of normals distribution:

$$p(\boldsymbol{\theta}_h|\boldsymbol{\pi}, \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}) = \sum_{k=1}^{K} \pi_k \phi(\boldsymbol{\theta}_h|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k). \tag{I.3}$$

The probability density of $\boldsymbol{\theta}_h$ is given as the weighted sum of K multivariate normal densities, each with mean $\boldsymbol{\mu}_k$ and covariance matrix $\boldsymbol{\Sigma}_k$. The weights $\boldsymbol{\pi} = (\pi_1, ..., \pi_K)$ sum to one. A straightforward interpretation of specification (I.3) is the following: One can think of K basic types of households, each of which characterized by a multivariate normal distribution of the coefficients $\boldsymbol{\theta}_h$ with mean $\boldsymbol{\mu}_k$ and covariances $\boldsymbol{\Sigma}_k$. π_k then is simply the probability that a given household is of type k.

Note at this point that the covariance matrices $\boldsymbol{\Sigma}_k$ are not restricted. That means my model specification is flexible enough to capture correlations in the brand preferences of the households. This kind of flexibility is important for the identification of inertia in households' umbrella brand choices. In order to illustrate this, let us assume for a moment that the covariance matrices $\boldsymbol{\Sigma}_k$ were restricted in a way which ruled out correlations among households' preferences for different brands. Assume further that in our data the preferences of households for brands under a certain umbrella were positively correlated. In that case my model would falsely attribute the tendency of households to stay with this umbrella to the existence of choice inertia with respect to umbrella brands. Not restricting the covariance matrices $\boldsymbol{\Sigma}_k$ controls for the possibility of correlated brand preferences, and thus allows to separate effects resulting from the existence of inertia in umbrella brand choice from effects resulting from correlated brand preferences.

To complete the econometric specification of model (I.1), I specify the error terms ϵ_{jt}^h to follow a type I extreme value distribution. The type I extreme value distribution is similar to the normal distribution. It is common to assume the error terms of discrete choice models to follow a type I extreme value distribution. The reason is that with this specification closed form expressions for households' choice probabilities can be derived.

■ **Bayesian estimation.** I use a Bayesian estimation approach to derive the posterior distribution of my model coefficients $\boldsymbol{\theta}_h = (\boldsymbol{\alpha}^h, \eta^h, \beta^h, \gamma^h)$. Bayesian estimation is conceptually very simple - it is essentially just a straightforward application of Bayes' rule: Conditional on prior information about the model coefficients $\boldsymbol{\theta}_h$, the posterior distribution of $\boldsymbol{\theta}_h$ given data $\mathbf{y} = (\mathbf{y}_1, ..., \mathbf{y}_H)$ (that is, $p(\boldsymbol{\theta}_h|\mathbf{y})$) is determined as

$$p(\boldsymbol{\theta}_h|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta}_h)p(\boldsymbol{\theta}_h)}{p(\mathbf{y})}. \tag{I.4}$$

$p(\mathbf{y}|\boldsymbol{\theta}_h)$ denotes the likelihood of the observed data $\mathbf{y}$, which is conditional on the model parameters $\boldsymbol{\theta}_h$, $p(\boldsymbol{\theta}_h)$ is the prior on the model parameters, and $p(\mathbf{y})$ is the unconditional probability to observe the data $\mathbf{y}$, which simply as normalizing constant.[7] In essence, equation (I.4) captures the whole estimation process. That is, given some prior information the posterior follows directly. What makes the Bayesian estimation approach computationally involved is just the computation of the right-hand side of (I.4). However, apart from the matter of specifying a sensible prior, this is only a technical matter. In the following I discuss how the prior $p(\boldsymbol{\theta}_h)$ is formed and how I compute the right-hand side of (I.4).

Specifying a prior directly on the distribution of the $\boldsymbol{\theta}_h$ would mean specifying prior values for the moments $\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$ of the normal components and the mixture probabilities $\boldsymbol{\pi}$. Problematic with this "direct" prior is that it treats every household in the same way regardless of the observations available for each household. This means that via households for which only a few observations are available the influence of the "direct" prior on the posterior might be quite strong.

An alternative approach is to make the prior on the household coefficients $\boldsymbol{\theta}_h$ household-specific. This can be achieved by the use of a two-stage prior. A two-stage prior specifies that for each household the $\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$ and $\boldsymbol{\pi}$ are themselves drawn from prior distributions with parameters $\mathbf{h}$. This way only the parameters $\mathbf{h}$ have to be directly specified for the whole sample of households, and the $\{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}$ and $\boldsymbol{\pi}$ are specific for every household and influenced by the number of observations available. Thus, the use of a two-stage prior allows more flexible adaptation to information in the data and thereby reduces the influence of prior information on the posterior.

In specifying a two-stage prior, I follow the approach of Rossi et al. (2005) and Dube et al. (2010) and specify a hierarchical prior with the mixture of normals (I.3) as first stage and a prior $\mathbf{h}$ on the parameters $\boldsymbol{\tau} \equiv \{\boldsymbol{\pi}, \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}\}$ of the mixture of normals

[7] It can be computed as $p(\mathbf{y}) = \int p(\mathbf{y}|\boldsymbol{\theta}_h)p(\boldsymbol{\theta}_h)\mathrm{d}\boldsymbol{\theta}_h$.

distribution as second stage. The posterior distribution of my model coefficients $\boldsymbol{\theta}_h$ is then given as

$$p(\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_H | \mathbf{y}_1, ..., \mathbf{y}_H, \mathbf{h}) \sim \left[\prod_h p(\mathbf{y}_h | \boldsymbol{\theta}_h) p(\boldsymbol{\theta}_h | \boldsymbol{\tau}) \right] \cdot p(\boldsymbol{\tau} | \mathbf{h}). \tag{I.5}$$

$\mathbf{y}_h$ denotes the data available for household h. The normalizing constant, which is not explicitly stated here, is the unconditional probability to observe the data $\mathbf{y}$ and is given as the integral of the product of the data likelihood and the prior density over the parameter space.

The posterior distribution cannot be expressed analytically. I therefore estimate the posterior distribution of my model parameters $\boldsymbol{\theta}_h$ by employing a modified version of the Markov-Chain-Monte-Carlo (MCMC) algorithm used by Dube et al. (2010) and described in detail in Rossi et al. (2005).

In contrast to the algorithm used by Dube et al. (2010), my modification of it allows for varying choice sets. This modification is important for my research purposes. The reason is that other than Dube et al. (2010), who are only interested in household inertia with respect to single products, in order to identify household inertia in umbrella brand choice I have to take into account a large number of products which are assembled under different umbrella brands. The drawback with this need to cover a large range of products is that only on rare occasions the whole set of products I consider coincides with the set of products a household can choose from at a given storeweek.[8] My modification of the algorithm of Dube et al. (2010) allows me to estimate choice model (I.1) also on storeweeks where only a subset of the whole set of products I consider is available and thus drastically increases estimation efficiency. Note that allowing for a varying choice set is only a technical matter which makes the algorithm which computes the posterior distribution more involved. There is no conceptual reason stemming from discrete choice theory which demands a constant choice set. Technical details on the prior specifications and the MCMC algorithm can be found in appendices A.1 and A.2.

I.3 Household Panel Data on Grocery Purchases

■ **Data description.** I estimate my choice model on household panel data on toothbrush purchases. The household panel data on toothbrush purchases is part

[8] Of course I could estimate my model only on storeweeks where households can choose from the whole set of products I consider. However, that would drastically reduce the number of observations available for estimation and thus render estimation inefficient.

of a large dataset collected by the IRI marketing institute. A detailed description of the dataset can be found in Bronnenberg et al. (2008). Household panel data is available for the years 2001 to 2005 for two metropolitan areas in the USA: Pittsfield, Massachusetts, and Eau Claire, Wisconsin. The data gives information about how many units of which product at what price were purchased by a certain household in a certain store in a certain week. It contains extensive information on each product, including information about its branding.

In particular, I use household panel data on toothbrush purchases in the seven largest grocery stores of the two metropolitan which was collected during the years 2001 to 2005. I use data from the toothbrush category for three reasons: First, the branding structure in this category is ideally suited for my research purposes - products in this category are assembled under several brands and these again under several umbrella brands. Second, households' purchasing behavior in this category fits the assumptions implicit in every discrete choice model very well. In particular, in most storeweeks households purchase exactly one unit of one product. Third, some umbrellas which are used in the toothbrush category are also used in the toothpaste category, which will allow me to analyze cross-category effects of umbrella branding. I supplement the panel data with store-level data to fill in information about the availability and the prices of products which were not purchased by any of the households in the panel in a certain storeweek. The store-level data contains information about all purchases made in a certain store in a certain week. For my estimations I use only households which were observed at least twice during the sampling period. That leaves me with 775 households. Table I.1 depicts summary statistics for these households.

I concentrate on the three largest umbrella brands in the market. In terms of purchases these umbrellas cover 67.7% of the market. Per umbrella I concentrate on products which account for at least 5% of purchases of this umbrella. This leaves me with 23 products. A product is defined as a toothbrush with a certain brand, a certain head size (compact vs. full) and a certain type of bristle (soft vs. medium). For the largest umbrella I observe seven products which are assembled under three brands, and for both the second- and the third-largest umbrella I observe eight products each which are assembled under four brands. Table I.2 describes the market for toothbrushes in the seven largest grocery stores in the metropolitan areas Eau Claire, Wisconsin, and Pittsfield, Massachusetts, for the years 2001 to 2005. The outside good is defined as any toothbrush sold in this market other than the 23

	Mean	Std. Dev.	Median	Min.	Max.
Households	775				
No. of shopping trips	5,263				
No. of shopping trips per hh.	6.8	6.8	5	2	101
No. of products purchased	6,930				
No. of purchases per household	10.5	10.9	7	2	139
Share of shopping trips where outside good was purchased	47.8%				
Grocery stores	7				
Share of purchases per store	14.3%	8.0%	10.2%	7.4%	27.8%

Table I.1: **Descriptive statistics (purchases in the toothbrush category).** The table gives descriptive statistics for the purchases the households in my sample made in the toothbrush category. My sample includes all households which were observed to shop toothbrushes at least twice in one of the seven largest grocery stores in the metropolitan areas Eau Claire, Wisconsin, and Pittsfield, Massachusetts, during the years 2001 to 2005.

toothbrushes considered. For confidentiality reasons I do not give umbrella/brand names in table I.2.

In 5,263 storeweeks, the 775 households I observe made discrete choices among the outside option and a certain subset of the 23 products I look at. The choice set households were confronted with varied from storeweek to storeweek. I make the assumption that every product in the choice set consumers were confronted with in a certain storeweek was sold at least once in this storeweek. This assumption implies that I can use store-level data to reconstruct the choice sets for every storeweek.

I observe every household in my panel for at least two storeweeks. In my estimations, for each household I use all observations apart from the first one. In doing so I circumvent the initial conditions problem. The costs are that I loose 15% of storeweek observations, but as the remaining number of storeweek observations is quite high that does not matter much in terms of estimation efficiency.

■ **Price variation.** In section I.2 I discussed that one key for identification of choice inertia and households' product preferences is variation in the prices of the observed products: For each household, the household's purchasing decisions given

Umbrella brand	**Brand**	**Average Price** ($, per blister)	**Purchases share**
A	A I	3.07	4.7%
	A II	2.89	2.3%
	A III	2.67	3.9%
B	B I	2.85	7.1%
	B II	3.18	1.7%
	B III	4.66	1.8%
	B IV	2.51	7.1%
C	C I	2.38	3.1%
	C II	2.39	1.1%
	C III	3.69	0.5%
	C IV	3.15	1.7%
Outside good		2.73	64.8%

Table I.2: **Brand structure of the market for toothbrushes.** The table displays the structure of the market for toothbrushes in the seven largest grocery stores in the metropolitan areas Eau Claire, Wisconsin and Pittsfield, Massachusetts, during the years 2001 to 2005. For confidentiality reasons the umbrella/brand names are not given.

different product price vectors reveal information about the relationships among this household's product preferences. In addition, purchasing decisions given changing price patterns reveal information about whether this household incurs switching costs.

Figure I.1 exemplarily depicts price patterns for two products under different umbrellas which were observed in the largest grocery store in the sample. As can be seen, over time a household which shops in this grocery store is confronted with quite different price vectors. The reactions of a household to the different price vectors allow identification of the household's product preferences. What is also evident is that after a deflection product prices often return to their initial levels, which leads to the occurrence of repeated price patterns. This repeated price patterns facilitate the identification of choice inertia. Note that the pricing patterns depicted in figure I.1 are representative for all products and all stores in my sample.

■ **Variation in households' choices.** Besides variation in prices, in order to identify choice inertia I need to observe variation in households' choices: Roughly put, choice inertia is identified when a household stays with a product it switched to in the previous period even when prices go back to their levels before the previous period. Table I.3 shows that there is indeed a lot of variation in households' choices.

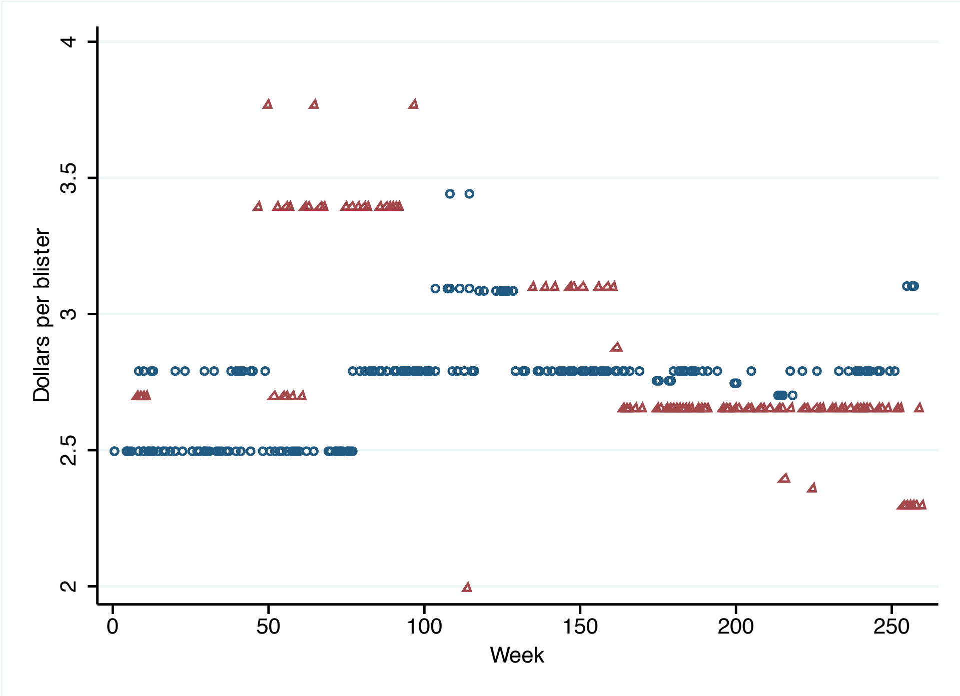

Figure I.1: **Exemplary pricing patterns.** The graph shows exemplarily pricing patterns for two products. The depicted pricing patterns are observed in the largest grocery store for two products under different umbrellas.

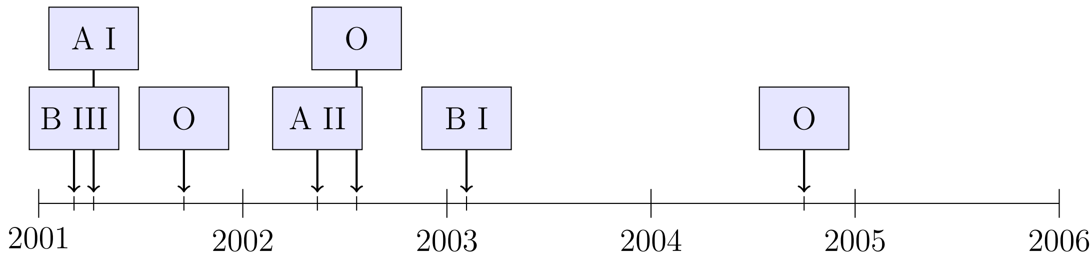

Figure I.2: **Exemplary purchasing pattern.** The timeline exemplarily shows the purchasing decisions a randomly chosen household made during the five year observation period. "O" denotes a purchase of the outside option. The capital letters mark umbrella brands, the latin numbers the brands assembled under the respective umbrella brands (compare table I.2).

On average, a household is observed to purchase toothbrushes on six shopping occasions. It changes the brand of the toothbrush roughly twice, and the umbrella brand roughly once. From these changes information about the existence of choice inertia can be inferred.

Figure I.2 exemplarily shows the purchasing behavior of a randomly chosen household. The depicted pattern is typical for the households in my sample. Clearly, there is variation in household's purchasing behavior: Over the five year observa-

	Mean	Std. Dev.	Median	Min.	Max.
Households	775				
No. of shopping trips per hh.	5.8	6.8	4	1	100
Different products per household	2.9	1.5	3	1	11
Different brands per household	2.6	1.2	2	1	8
Different umbrellas per household	2.1	0.8	2	1	4

Table I.3: **Descriptive statistics (product, brand and umbrella brand purchases).** The table gives descriptive statistics about the number of different products, brands and umbrella brands purchased by the households in my sample. The sample includes all households which were observed at least twice to shop toothbrushes in one of the seven largest grocery stores in the metropolitan areas Eau Claire, Wisconsin, and Pittsfield, Massachusetts, during the years 2001 to 2005.

tion period the household changes both among umbrellas and among the brands assembled under an umbrella. Together with the observed variation in prices (as exemplarily depicted in figure I.1), these changes allow the identification of choice inertia, both with respect to brands and umbrella brands.

■ **Indication for the existence of inertia in umbrella brand choice.** Table I.4 displays the results of a naive logit discrete choice estimation on all the households in my sample. The results show that if we control for product prices households actually seem to exhibit inertia both in their brand and umbrella brand choices, and that the effect of inertia is quite strong. However, as the results in table I.4 are not derived from the observation of one household but rather reflect the average behavior of all households in my sample, the observed inertia in households' choices cannot be interpreted to be structural. Indeed, it might simply be rooted in differences between households with respect to their brand and umbrella brand preferences or in correlations among product tastes. Thus, the results in table I.4 give only a first indication that structural choice inertia with respect to umbrella brands might exist. In order to separate structural choice inertia from preference heterogeneity and correlated product tastes I have to estimate a model which flexibly accommodates different forms of household heterogeneity and preference correlations.

Covariates in household's utility fct.	**Coefficient estimate**	**Standard Error**
Product price ($)	-.533***	.049
Brand change	-.778***	.073
Umbrella change	-.562***	.063
Nbr. of observations	68,298	
Nbr. of households	775	

Table I.4: **Results of naive logit estimation.** The table gives results of a naive logit estimation. The logit discrete choice model assumes the utility u^h_{jt} a household h derives from the choice of product j at storeweek t to be given as $u^h_{jt} = \eta p_{jt} + \beta \mathbb{1}_{B(j)\neq B_{t-1}} + \gamma \mathbb{1}_{U(j)\neq U_{t-1}} + \epsilon^h_{jt}$. If the outside option is chosen the household derives utility $u^h_{0t} = v_0 + \epsilon^h_{jt}$. The ϵ^h_{jt} are assumed to be iid type I extreme value distributed. The value v_0 of the outside option is not significantly different from zero, for which reason I do not report its estimate here. The estimates are based on data on choices 775 households made on 5,263 shopping trips. Significance niveaus are reported by stars: ***: 1%, **: 5%, *: 10%.

I.4 Main Results

I estimated choice model (I.1) on the household panel data described in chapter I.3 using the mixture of normals specification (I.3). With this specification model (I.1) flexibly accommodates different forms of household heterogeneity and preference correlations. I chose K, the number of normal components, to equal 5, and I chose very diffuse and thus non-informative priors on the parameters $\{\boldsymbol{\pi}, \{\mu_k, \boldsymbol{\Sigma}_k\}\}$ of the mixture of normals distribution. The exact prior specification is given in appendix A.1, and I will discuss the choice of this specification and the choice of the number of normal components in section I.5. The upper-left graph in figure I.3 shows the key result of my estimation: Most of the probability mass of the posterior distribution of coefficient γ^h is on negative values. That is, most households exhibit structural inertia in umbrella brand choice.

Besides the posterior distribution of γ^h figure I.3 depicts the posterior distribution of coefficient β^h, which captures inertia in brand choices, and the posterior distribution of the price coefficient η^h. Most of the probability mass of the posterior distribution of coefficient β^h is on negative values, which means that besides inertia in umbrella brand choice the majority of households also exhibits inertia in brand choice. As is to be expected, the probability mass of the posterior distribution of the price coefficient η^h is nearly entirely on negative values. The small part of the probability mass which is on positive values is an artifact of my mixture of normals specification.

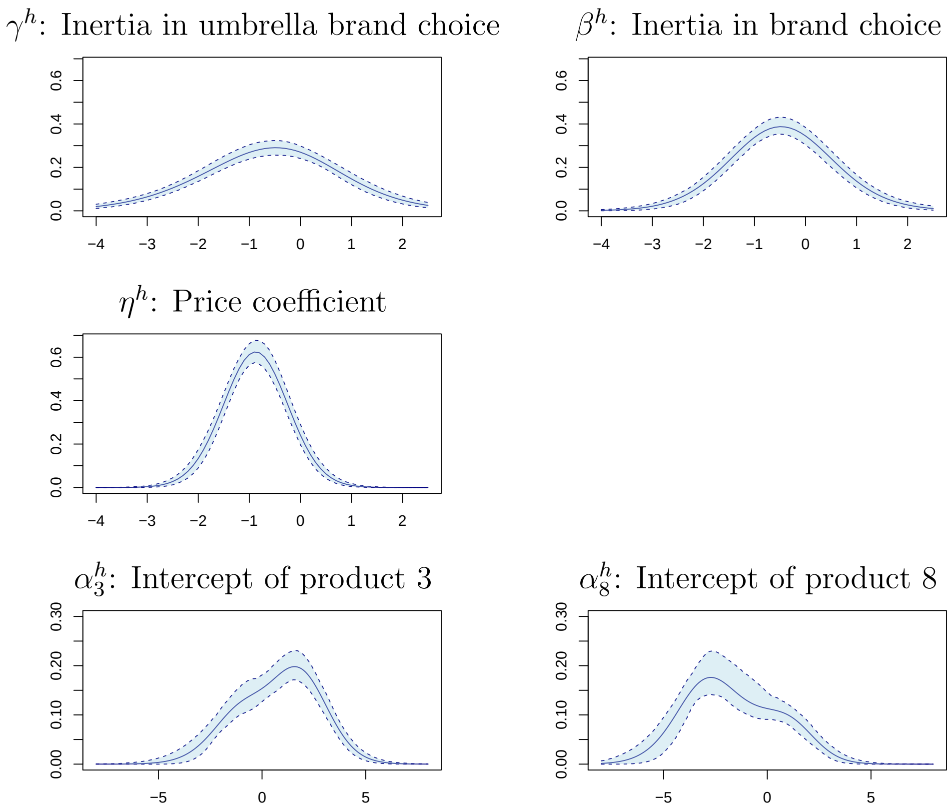

Figure I.3: **Main estimation results (posterior marginal distributions).** The graphs depict the pointwise posterior means and the 95% credibility regions of the marginal densities of the price coefficient, the coefficient capturing inertia in brand choice, and the coefficient capturing inertia in umbrella brand choice. In addition, in order to demonstrate the need for a flexible preference specification, also the posterior distributions of the product-specific intercepts α_6^h and α_8^h are shown. The results are based on 5,263 purchasing observations of 775 households, and were derived by estimating choice model (I.1) given the non-informative five component prior specification detailed in appendix A.1.

Table I.5 gives some summary statistics on the posterior distributions of γ^h, β^h and η^h. For the mean household the value of γ^h is of the same order of magnitude as the value of η^h. That is, for the mean household inertia in umbrella brand choice has a similar impact on its purchase decision as a price change of one dollar. Or, in more illustrative terms: Based on the figures in table I.5, conditional on a brand change the mean household is indifferent between changing to a brand under the same umbrella as the previously purchased brand and changing to a brand under another umbrella which is by \$0.69 cheaper. (The simplifying assumption here is

	Price coefficient (η^h)	**Inertia in brand choice (β^h)**	**Inertia in umbrella brand choice (γ^h)**
Mean	-0.864	-0.452	-0.594
Std. error of mean	0.049	0.103	0.099

Table I.5: **Summary statistics of posterior marginal distributions.** The table gives summary statistics on the posterior marginal distributions of the price coefficient η^h, the coefficient β^h which captures inertia in brand choice, and the coefficient γ^h which captures inertia in umbrella brand choice (see figure I.3). In the limit of infinite observations these numbers would equal traditional frequentist maximum likelihood estimates of model (I.1).

that a priori there is no difference between brands and prices.) Relative to the mean product price of \$2.73 a price reduction of \$0.69 equals a price decrease by 25%.[9]

For completeness and in order to demonstrate the necessity of a flexible preference specification, figure I.3 exemplarily depicts the posterior distributions of the product-specific intercepts α_6^h and α_8^h. The product-specific intercepts $\boldsymbol{\alpha}$ capture households' product preferences. As mentioned in section I.2, for the identification of household inertia in brand and umbrella brand choice - that is, for the identification of the coefficients γ^h and β^h - it is crucial to capture heterogeneity in households' product preferences. As can be seen from the distributions of the product-specific intercepts α_6^h and α_8^h, standard distributional assumptions (like for example that of a simple normal distribution) are obviously not suited to sufficiently capture heterogeneity in households' product preferences. Thus, it is important to use a distributional specification which is able to flexibly accommodate various forms of heterogeneity.

I.5 Robustness of Estimation Results

I derived the results given in section I.4 using choice model (I.1) with the non-informative prior specification given in appendix A.1. In the following I will discuss the robustness of these estimation results. The discussion will involve comparisons of my model of choice to other possible models. In the Bayesian framework models

[9]The effect of inertia in brand choice, which is captured by coefficient β^h, is on average equivalent to the effect of a price change of 19%. This figures are of the same order of magnitude as numbers from Dube et al. (2010), who find that switching costs with respect to single products on average amount to 12% of mean product price in the margarine product category, respectively 21% of mean product price in the orange juice product category.

Number of normal components:	
1 component	-6324.615
3 components	-6121.99
5 components (BM)	-6064.658
10 components	-6004.293
Concentration parameter:	
$a = 0.1$ (BM)	-6064.658
$a = 1.5$	-6057.338
Model specification; addition of dynamic terms:	
$u_{jt}^h = \alpha_j^h + \eta^h p_{jt} + \epsilon_{jt}^h$	-6245.72
$u_{jt}^h = \alpha_j^h + \eta^h p_{jt} + \beta^h \mathbb{1}_{B(j) \neq B_{t-1}} + \epsilon_{jt}^h$	-6247.769
$u_{jt}^h = \alpha_j^h + \eta^h p_{jt} + \beta^h \mathbb{1}_{B(j) \neq B_{t-1}} + \gamma^h \mathbb{1}_{U(j) \neq U_{t-1}} + \epsilon_{jt}^h$ (BM)	-6064.658
Store-specific intercepts and advertising controls:	
Addition of store-specific intercepts	-5950.316
Addition of advertising controls	-5916.405

Table I.6: **Marginal log-likelihoods for different model specifications.** The table displays marginal log likelihoods of models which differ either with regard to their prior specifications or the model specification itself. Apart from the deviations mentioned explicitly in the table, the specification of the models equals (I.1), and the prior specifications are those given in appendix A.1. The base model with which the main results from section I.4 were derived is marked by “BM”.

can be compared based on posterior model probabilities. The posterior probability of a model is simply the probability that this model is “true” given the data and the prior information at hand. As detailed in Rossi et al. (2005) and Dube et al. (2010), model choice on basis of posterior model probabilities is consistent, meaning that with increasing sample size the probability of choosing the true model tends to one.

Under the assumption of equal prior model probabilities model comparison on basis of posterior model probabilities is equivalent to model comparison on basis of model marginal likelihoods. Following the treatment in Dube et al. (2010), technical details regarding the equivalence of model comparison on basis of posterior model probabilities and on basis of model marginal likelihoods and the computation of marginal model likelihoods are given in appendix A.3. Important to keep in mind for the following is that (under the assumption of equal prior model probabilities) model choice on basis of model marginal likelihoods is consistent in the sense that with increasing sample size the probability of choosing the true model tends to one.

Finally, note that model marginal likelihoods automatically adjust for the parameter dimensions of models. That is, there is no positive discrimination of models which are large in terms of their parameter dimensions.

Robustness against misspecification of preferences

One concern with my estimation results from section I.4 might be that the coefficients γ^h in model (I.1) do not pick up dynamic effects, but rather simply capture some part of household heterogeneity or correlation among product tastes. If the umbrella brand inertia coefficients γ^h simply captured some part of household heterogeneity or taste correlation, then their addition to model (I.1) should not higher the posterior model probability: My econometric specification does not put any restrictions on the correlations among the α_j^h, and the mixture of normals specification is able to very flexibly accommodate various forms of distributions. That is, the α_j^h should fully capture both arbitrary correlations among product tastes and heterogeneity among households. Thus, if the γ^h simply picked up some part of household heterogeneity or taste correlation, their addition would mean overfitting my model. As shown by Dube et al. (2010), overfitting decreases the posterior probability of a model. Table I.6 shows that on the contrary the addition of the umbrella brand inertia coefficient γ^h strongly increases the posterior model probability. This indicates that the umbrella brand switching cost coefficient γ^h does not simply pick up heterogeneity in or correlations among households' preferences but actually captures structural inertia in umbrella brand choice.

Robustness under different prior specifications

■ **Prior on the number of normal components.** In theory mixtures of normals distributions can be used to approximate any kind of continuous distribution with full support. The quality of the approximation depends on the number of normal components used. In general, the higher the number of components the better the approximation. However, with an increasing number of normal components the problem of overfitting arises.[10]

[10] Overfitting means that the estimation results do not pick up general patterns in households' preferences and choice behavior but capture noisy behavior. As an extreme example, this would be the case if the number of components was equal to the number of households. Then every component would simply pick up a specific households' behavior, and the model would have no explanatory and predictive power at all.

For the derivation of the results given in section I.4 I used five normal components. With five normal components overfitting does not pose a problem: Given the interpretation that each component of the multivariate normal distribution represents one basic type of household, and given that I have 775 households in my sample, each component should capture general characteristics of households and should not pick up noisy behavior. On the other side, this number of components seems to be sufficient to fully capture heterogeneity in households' preferences: Table I.6 displays the marginal log-likelihoods for models with one, three, five and ten normal components. The marginal log-likelihood, which is equivalent to the posterior model probability, strongly increases from the model with one component to that with five components. From the model with five normal components to that with ten normal components there is only a slight increase in the marginal log-likelihood. This pattern suggests that the basic heterogeneity pattern is picked up when five components have been added to my model, and that the addition of further components does not add much to the understanding of the general preference structure of the households.

A look at figure A.1, which can be found in the appendix and which exemplarily depicts posterior distributions for models with one, five and ten components, confirms the aforementioned. The posteriors of the five- and the ten-component model clearly deviate from those of the one-component model. This emphasizes the need for a heterogeneity specification more flexible than that of a simple normal distribution. The posteriors of the ten-component model deviate in detail from those of the five-component model but exhibit the same main characteristics. In particular, addition of normal components beyond the fifth one is of insignificant effect on the posterior distributions of γ^h, β^h and η^h.

■ **Prior on composition of average household.** The prior on how a household is composed from the basic types of households (that is, the normal components) determines in what way the basic household types are combined to produce a given household: In expectation, a given household might either resemble one of the basic types or a balanced mixture of all the basic types. I derived the results in section I.4 under the prior assumption that a given household in expectation resembles one of the basic types of households. Concretely, I set the concentration parameter a of the symmetric Dirichlet distribution, which determines the composition of an average household, to 0.1. For details, see appendix A.1. I compared the posterior distributions resulting from this prior assumption about the composition of an average household to that that an average household is a balanced mixture of all basic types.

In more technical terms, I changed the concentration parameter a from 0.1 to 1.5. It showed that different prior assumptions about how an average household is composed from the basic household types do not have significant effects on the posterior coefficient distributions. Figure A.2 in the appendix illustrates this by comparing the posterior distributions for the different concentration parameters. The small influence of the concentration parameter is mirrored in the fact that the marginal log likelihood does not significantly change when the concentration parameter is varied (see table I.6).

■ **Prior on the distribution of household coefficients.** In order to ensure that the priors on the mean and the variance of the distribution of my model coefficients are of negligible effect on my qualitative results I chose them to be neutral and very diffuse. In particular, I used a prior specification such that in expectation for every coefficient and every component the prior mean of the coefficient distribution is zero and the prior variance is five. A prior coefficient mean of zero implies that a priori my model is neutral with respect to the direction of the effects the coefficients are supposed to capture. The prior coefficient variance of five becomes meaningful in relation to my estimation results: The average effect sizes seem to lie in the range between around 0.5 and 1. Given these effect sizes prior coefficient variances of five mean that my model is a priori quite non-informative with respect to the exact location of the coefficients.

For the price coefficient η^h, the brand choice inertia coefficient β^h, and the umbrella brand choice inertia coefficient γ^h figure I.4 contrasts the (expected) prior coefficient distribution to the posterior coefficient distributions. The posterior coefficient distributions strongly deviate from the prior coefficient distributions, both with regard to location and spread. This demonstrates that my results are to a large extent driven by information from the data and not by prior information.

Comparison of the prior specification which leads to prior coefficient variances of five and which I used to derive my main results in section I.4 to a tighter prior specification which leads to prior coefficient variances of 2.5 shows that the tighter prior has no significant effect on the location of the posterior distributions. It has, however, an effect on the spread of the posterior distributions: The tighter the prior specification, the less wide-spread the posterior distributions. Although this effect is not very strong I work with a very diffuse prior specification in order to render the effect of prior information on my qualitative results as marginal as possible. Figure

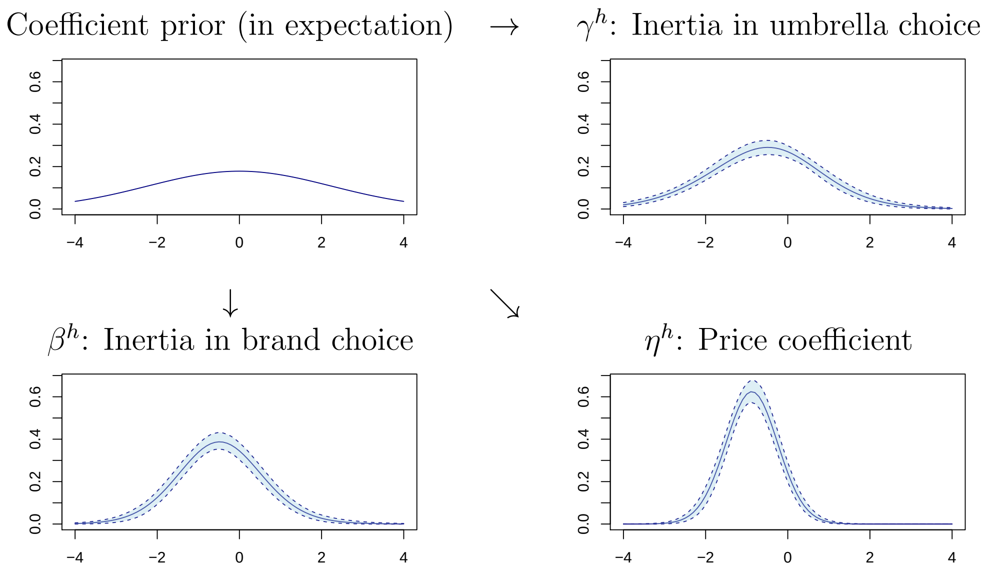

Figure I.4: **Comparison of prior coefficient distribution and posterior coefficient distributions.** The upper-left graph shows the common prior marginal distribution (in expectation) of the price coefficient, the brand choice inertia coefficient and the umbrella brand choice inertia coefficient. The other graphs depict the posterior marginal distributions of these coefficients. The results were derived with the base prior specification. The base prior specification is detailed in appendix A.1. The results are based on 5,263 purchasing observations of 775 households.

A.3 from the appendix exemplarily displays posteriors for the diffuse and the tight prior specification.

Robustness against marketing measures

The technical discussions above showed that the results given in section I.4 are robust under different prior specifications, and that the choice inertia coefficients β^h and γ^h in model (I.1) actually pick up structural state dependence in brand respectively umbrella brand choice. In this section I turn to a discussion of my results from section I.4 which focuses more on the economic mechanics my model tries to capture. In particular, I assess the key assumption which allows identification of my model coefficients: This key assumption is that there are no influences on households' choice behavior which are systematically connected to the explanatory variables of model (I.1) but unobserved by me as econometrician.

In particular, as pointed out by Dube et al. (2010), in order for my model to identify structural choice inertia it is necessary that the price coefficient estimates are unbiased. The reason is that if the price coefficient (which measures a household's sensitivity to price changes) is not determined correctly, my model might interpret household behavior actually induced by price variations to be caused by inertia in umbrella brand choice. As a simple example, if the estimate of the price coefficient was biased towards zero and a household purchased a certain brand or umbrella repeatedly simply due to low prices, my model could attribute this behavior spuriously to the existence of switching costs.

Price coefficient estimates are possibly biased when there are measures which are controlled by the supply side, which influence a household's inclination to purchase a product, and which are accounted for when the product price is set. Such measures can be taken on or above store-level. For example, a store whose strategy is to increase households' willingness to spend money by creating a pleasant shopping experience (example through ample space and an appealing interior design) might in turn demand product prices which are above average. Or, a producer who launches a large marketing campaign in order to make households aware of his brand might in turn demand a price premium for products under the advertised brand. In the following I turn to a discussion of how marketing measures on or above store-level might possibly affect the estimation results presented in section I.4.

■ **Marketing measures on store-level: Store characteristics.** The households in my sample are observed to make purchases in seven grocery stores. Things like the layout of these stores, the shelf design or the style of the interior might influence households' willingness to spend money on the products offered in different ways, and the way in which households' purchasing decisions are influenced might be accounted for when final product prices are set. If there actually is a systematic but unobserved correlation between the characteristics of a store and the prices of its products, then this would potentially bias the estimates of the price coefficients η^h, which in turn could lead to spurious identification of choice inertia.

To account for influences of the characteristics of a store on the purchasing decisions of households I introduce store-specific intercepts into my model, which make the utilities households derive from product purchases store-dependent. In doing so, I control for possible systematic correlations between store characteristics and households' purchasing decisions.

The prior settings for the store-specific intercepts are the same as for the price and the switching cost coefficients. Figure A.4 compares posterior marginal distributions derived from the model with store-specific intercepts to that derived from my base model (I.1). Comparison of the posterior marginal distributions derived from the model with store-specific intercepts to that derived from my base model (I.1) shows that inclusion of store-specific intercepts does not alter the marginal posterior distributions in a significant way. Thus, store characteristics seem not to systematically influence households' choice behavior in a way which biases my main estimation results given in section I.4.

■ **Marketing measures on store-level: In-store advertisement and display.** With the introduction of store-specific intercepts I control for time-invariant store characteristics which might be systematically correlated with households' purchasing decisions. Time-varying measures on the store-level which potentially influence households' purchasing decisions are in-store advertisement and in-store display of products. In my data I have available information on whether a certain product was advertised or put on display in a certain store in a certain week. Basically, data on in-store advertisement and in-store display is available separately. However, as in the toothbrush category display occasions are very rare (they occur only in around three percent of the storeweeks), for efficiency reasons I combine the data on in-store advertisement and in-store display. In modeling terms, I introduce a flag f^h in the utility a household derives from a certain product whenever this product is either advertised or put on display in a certain store in a certain week. In doing so I am able to control for the effects of in-store advertising measures on households' purchasing behavior.

The prior settings for the in-store advertisement flag are the same as for the price and the switching cost coefficients. Figure A.4 displays the marginal posterior distributions of the price coefficients η^h and the umbrella brand choice coefficients γ^h. As with store-characteristics, inclusion of controls for in-store advertising and display does not lead to posterior distributions which are significantly different from those derived from my base model (I.1). Therefore, in-store advertisement and display do not affect households' decisions in a way which biases my main estimation results from section I.4.

■ **Marketing measures above store-level: Mass media advertising.** Households' purchasing decisions might not only be influenced by prices and by marketing measures on the store-level, but also by product or brand advertising via channels like newspaper ads or TV commercials. Erdem et al. (2008) conducted a study in which they analyzed the effects of advertisement exposure on households' willingness to pay for products from four categories, among them toothbrushes. They exclusively had available both scanner data and data from telemeters, which measured each household's specific exposure to TV commercials. They found that in general a household's willingness to pay for a certain product is increased if this household is exposed to advertisement of this product.

In terms of model (I.1) this can be interpreted as if exposure to product advertisement constitutes a positive shock to the (expected) utility the household derives from the purchase of the advertised product. In slightly more technical terms that means that ϵ_{jt}^{h} is likely to be large if household h at shopping trip t has recently been exposed to advertisement of product j. As I do not have available data on the specific advertisement exposure of the households in my sample these shocks are unobserved to me as econometrician. The question now is whether unobserved shocks like these represent a serious obstacle to the identification of choice inertia.

Note first that as long as advertising has no dynamic effects on households' purchasing behavior its presence is unproblematic for the identification of choice inertia. That is, as long as the effects of advertising vary only over the household or the product dimension (for example, because households are exposed to different types of mass media with different types of brands advertised in it, or because the advertising intensities for different brands vary), then this variation is captured by the intercepts α_j^h, which are household- and product-specific.

In contrast, problems would arise if a household's decision to purchase a certain brand systematically coincided with exposure to advertisement of this brand. In this case my model would spuriously attribute repeated purchases of a brand or umbrella brand to the existence of choice inertia. However, for two reasons I do not think that there is a systematic correlation between households' purchases of certain brands and timely exposure to advertisement of these brands. First, I argue that it is unlikely that households are induced to go on a shopping trip by the exposure to advertisement of a toothbrush. Instead, as toothbrushes are goods with a relatively low perishability, it suggests itself that from time to time they enter a household's shopping cart when it is shopping goods with a higher perishability. This reasoning is confirmed by figures given by Bronnenberg et al. (2008), which show that toothbrushes in general are bought besides other grocery products, and that toothbrushes, like for example razors and blades, have relatively long purchase cycles. Second, as reported by Erdem et al. (2008), the variation in advertising exposure is quite high. Given (with respect to toothbrush purchases) exogenous shopping trips, that makes it quite unlikely that a household's shopping trips are systematically correlated with exposure to advertisement of only one certain brand or umbrella.

Put together, a household's purchasing decisions might of course be influenced by its exposure to advertising of certain brands or umbrellas. This would bias my main results from section I.4 if the timing of a household's shopping trips was correlated with its exposure to advertising of certain brands or umbrellas. However, as in general a household's decision to go on a shopping trip should be unrelated to its exposure to the advertisement of a toothbrush, and as according to Erdem et al. (2008) the variation in a household's advertising exposure is very high, correlation between a household's exposure to the advertising of certain brands or umbrellas and its shopping trips is quite unlikely. Thus, as with marketing measures on the store-level, the observed choice behavior of households should not systematically be influenced by large-scale advertising measures.

I.6 Analysis of Behavioral Mechanism

Former studies, like Keane (1997), Erdem (1998), Seetharaman et al. (1999) and Dube et al. (2010), have shown that there is inertia in households' product choices. Simply put, households have a tendency to stay with the product they previously purchased. So far, I have demonstrated that in addition to inertia in product choice there is inertia in umbrella brand choice. That is, if households change away from the product they previously purchased, they seem to have a tendency to change to a product which is under the same umbrella brand as the previously purchased product.

In order to assess the economic implications of the finding of household inertia in umbrella brand choice it is necessary to get further insights into the mechanisms behind this phenomenon. Dube et al. (2010) proposed three possible explanations for household inertia in product choice: Search costs, learning about the quality of products, and structural (respectively psychological) switching costs. They find that the patterns in their data can most likely be explained by the existence of structural switching costs. The length of my household panel (I observe a major part of the households over the whole five year observation period) allows me to test whether my finding of inertia in umbrella brand choice also is rooted in the existence of psychological switching costs, or whether it can be explained by search or learning costs.

The basic idea is to split the whole observation period into two subperiods, an early one and a late one. For the late subperiod I then simply estimate choice model

	Mean	Std. Dev.	Median	Min.	Max.
Households	242				
No. of shopping trips per hh. ...					
...in years 2001 and 2002	6.1	6.3	5	2	78
...in years 2003 to 2005	5.5	6.4	4	1	71
No. of hhs. which in years 2001-2002 ...					
...purchased two umbrella brands	199				
...purchased three umbrella brands	43				
No. of hhs. which visited...					
...no new store in years 2003-2005	187				
...one new store in years 2003-2005	48				
...two new stores in years 2003-2005	7				

Table I.7: **Descriptive statistics for the subsample of experienced households.** The table shows descriptive statistics for the subsample of experienced households. The subsample includes all households which purchased at least two of the three umbrella brands in the sample in the years 2001 and 2002 and who were observed at least once in the years 2003-2005. It consists of 242 households. The figures depicted relate to shopping trips made in the seven largest grocery stores in the metropolitan areas Eau Claire, Wisconsin, and Pittsfield, Massachusetts.

(I.1) on a panel of experienced households, where experienced means that during the early subperiod these households have purchased products under at least two of the three umbrella brands considered, and that they have shopped in the same set of grocery stores during the early and the late subperiod. I argue that in the late subperiod these households have fully formed preferences regarding the different umbrella brands, and that they also know about the set of products available and where to find each product in each store. That is, for experienced households search and learning costs should play no role during the late subperiod . In consequence, if I estimate choice model (I.1) on the subsample of experienced households for the late subperiod only, and if the choice inertia I observed for the full sample of households was solely caused by search or learning costs, then I should no longer find household inertia in umbrella brand choice. If, however, I still observe inertia in umbrella brand choice for the subsample of experienced households, this will be evidence for the existence of structural (respectively psychological) switching costs.

Concretely, I split my data into an early subperiod covering the years 2001 and 2002, and a late subperiod covering the years 2003 to 2005. Into my subsample of experienced households I take all households who purchased at least two of the three

	Full sample of hhs.	Selected sample of experienced hhs.
γ^h: Inertia in umbrella choice	-0.594 (0.099)	-0.343 (0.151)
β^h: Inertia in brand choice	-0.452 (0.103)	-0.090 (0.179)
η^h: Price coefficient	-0.864 (0.049)	-1.095 (0.089)

Table I.8: **Comparison of the estimation results for the full sample and for the subsample of experienced households.** The subsample of experienced households contains 242 households and is a subset of the full sample of households (775 households). Depicted are the means of the marginal posterior coefficient distributions. The standard error of the means are given in parentheses. Both the results for the subsample and the full sample were derived from shopping trips in the seven largest grocery stores in the metropolitan areas Eau Claire, Wisconsin, and Pittsfield, Massachusetts. The results for the full sample of households were derived from data on shopping trips in the years 2001-2005, those for the subsample of experienced households from data on shopping trips in the years 2003-2005.

umbrella brands in the sample during the early subperiod, and who at least once are observed to make a purchase during the late subperiod. This leaves me with 242 households. As shown in table I.7, 199 households have purchased two umbrella brands during the early time segment. The remaining 43 households have purchased all three umbrella brands. 187 of the 242 households visited the same set of grocery stores during the late subperiod and the early subperiod. The remaining 55 panelists went to at most two new stores during the second subperiod (48 tried one new store, 7 tried two new stores). Based on these facts I make the assumption that from the year 2003 on the households in the subsample are experienced with respect to both the umbrella brands and the grocery stores. Thus, during the second subperiod if at all these households should only slightly be affected by search or learning costs.

I estimated choice model (I.1) on the subsample of experienced households. Compared to the results derived with the full sample of households some estimation efficiency is lost, and the prior is of stronger influence. As can be seen in figure A.5, this shows itself by larger credibility regions and more widespread distributions. However, the credibility regions still remain quite narrow, and information from the data clearly overwhelms the prior information. Thus, despite a considerably smaller dataset the results remain meaningful. Important to note is that also for the subsample of experienced households I find inertia in umbrella brand choice: Most of the probability mass of the coefficient γ^h is on negative values. Also, as can be seen from table I.8, the effect of inertia in umbrella brand choice is of similar size for both the subsample of experienced households and the full sample.[11] As the households in my subsample are experienced with respect to umbrella brands and grocery stores when they enter my estimation sample, this finding excludes both search and learning costs as possible explanations for household inertia in umbrella choice. Instead, it seems that inertia in umbrella brand choice is caused by structural (respectively psychological) switching costs.[12]

I.7 Cross-Category Inertia in Umbrella Brand Choice

Above I have used data on toothbrush purchases to show that households incur switching costs when they change umbrella brands. More specifically, I have shown that apart from the costs from a change of toothbrush products households incur additional costs when they switch from one family of toothbrush products marked by a certain umbrella brand to another family marked by another umbrella brand. That is, so far I have put forward evidence for the existence of umbrella brand related switching costs *within* a product category.

It is quite common that firms use umbrella brands to mark products from *several* product categories as being from the same product family. Erdem (1998) has already demonstrated that a link established by an umbrella brand connects households' quality perceptions of products from different product categories.[13] A related but open question is whether a cross-category link of products by an umbrella brand in addition leads to structural inertia in umbrella brand choice across categories.

[11]Actually, the effect of inertia in umbrella brand choice seems to be slightly less pronounced for the subsample of experienced households than for the full sample. This hints to the fact that to some smaller extent also search or learning costs might play a role.

[12]This finding is in line with that of Dube et al. (2010), who showed that inertia in product choice can most likely be explained by the existence of structural switching costs.

[13]Theoretically model (I.1) in combination with specification (I.3) is suited to replicate Erdem (1998)'s study, that is to analyze whether households' quality perceptions of products which are assembled under an umbrella brand are connected: If households' quality perceptions of products under a certain umbrella were connected, then one should observe that for one normal component (respectively one base type of household) the preference coefficients α_j are correlated for all products j under this umbrella. However, in practice the hybrid MCMC algorithm used to estimate the posterior does not necessarily identify the single normal components, as during the iterations switches between component labels might occur (for details compare Rossi et al., 2005). This behavior of the MCMC algorithm hinders the researcher to make statements about isolated normal components (respectively the base household types). However, as the households' preference distribution as a whole is identified (and thus controlled for), the possible occurrence of label switching does not put any restrictions on statements about the existence of household inertia.

	Mean	SD	Median	Min.	Max.
Households	775				
No. of households (hhs.) which...					
... bought toothpaste (tp.)	763				
... bought tp. under umbrella A	614				
No. of shopping trips with tp. purchase	13,956				
No. of shopping trips with tp. purchase per hh.	28.25	19.65	24	0	111
No. of toothpastes purchased	19,029				
No. of toothpastes purchased per hh.	42.89	32.40	35	0	198
Market share of umbrella brand A in toothpaste category	41.01%				
Percentage of purchase observations in toothbrush category with previous purchase of toothpaste under umbrella A	27%				

Table I.9: **Descriptive statistics for households' purchases in the toothpaste category.** The table gives descriptive statistics for the purchases households made in the toothpaste category during the observation period (that is, the years 2001 to 2005).

That is: When controlling for preference heterogeneity, is a household which in the previous period purchased a product under a certain umbrella in a certain category ceteris paribus more likely to buy a product under the same umbrella when it makes a purchase in another category?

The dataset I use contains household panel data from several product categories. Besides data from the toothbrush category I have available data from the toothpaste category. One of the umbrella brands in my toothbrush sample, umbrella brand A, is also present in the toothpaste category, with a market share of around 40%.[14] The occurrence of umbrella brand A both in the toothbrush and the toothpaste category offers the possibility to study whether inertia in umbrella brand choice can be observed not only within but also across product categories.

Table I.9 describes the purchasing behavior of households in the toothpaste category. Nearly all households purchased a toothpaste at least once during the five year observation period. Around 80% of the households at least once purchased a toothpaste which is under umbrella brand A. Compared to the numbers about

[14]There are also toothpastes which are sold under umbrella brand B. However, the market share of umbrella brand B in the toothpaste category is marginal (only about 0.03%).

toothbrush purchases given in table I.1 it is also evident that households purchase toothpastes more frequently than toothbrushes: They purchase toothpastes around four times as often as toothbrushes. These facts - that households can be expected to have purchased a toothpaste before they purchase a toothbrush, and that in around 40% of the cases this toothpaste can be expected to be from umbrella A - allow me to analyze whether previous purchases in the toothpaste category have an effect on purchases in the toothbrush category.

If inertia in umbrella brand choice across categories exists, then a previous purchase of a toothpaste under umbrella brand A should have a different effect on a household's decision for toothbrushes which are also under umbrella A than on its decision for toothbrushes which are under different umbrellas: Ceteris paribus, the household's inclination to purchase a toothbrush which is also under umbrella brand A should be higher than his inclination to purchase a toothbrush which is under another umbrella brand. In terms of model (I.1) this means that a previous purchase of a toothpaste under umbrella A should have a positive impact on the household's (expected) utility from a toothbrush which is under umbrella A, whereas it should have a negative impact on the expected utility from a toothbrush which is under another umbrella brand.

I model the effect of a previous purchase of a toothpaste under umbrella A on a household's decision among toothbrush products by including a covariate into model (I.1) which equals one if a household purchased a toothpaste product under umbrella A in the previous period, and zero otherwise. As I expect the effect of a previous-period purchase of a toothpaste under umbrella A to be different for toothbrushes also under umbrella A and for toothbrushes under other umbrella brands, I allow the coefficient on this covariate to be different for toothbrush products which are under umbrella A and which are under different umbrella brands. I estimate this model using the prior specifications detailed in appendix A.1.

Figure I.5 and table I.10 show the estimation results. The first thing to note is that when accounting for toothpaste purchases my main estimation results change only to some small extent. This is reassuring, as it means that the estimation results described in section I.4 are not strongly biased by not accounting for cross-category inertia effects, and it also does not come as a surprise: For less than one-third of the observed toothbrush purchases in the previous period a toothpaste under umbrella A had been purchased. Thus, in order to strongly bias the estimation results the

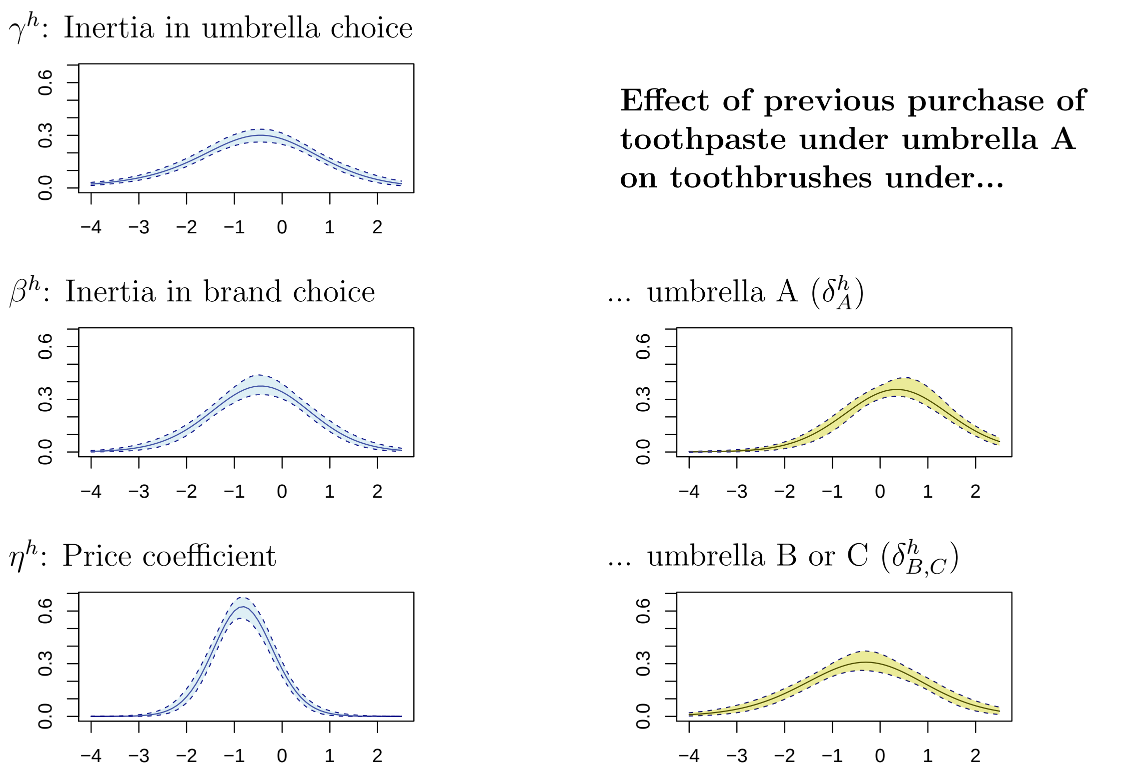

Figure I.5: **Posterior marginal distributions for model which accounts for cross-category inertia in umbrella brand choice.** The choice model is based on model (I.1) but includes a covariate indicating whether the household previously purchased a toothpaste under umbrella A. The coefficient on this covariate is allowed to be different for utilities from toothbrush products also under umbrella A (δ^h_A) and for products under umbrellas B and C ($\delta^h_{B,C}$). All graphs depict pointwise posterior means and 95% credibility regions of marginal posterior densities. The results are based on 5,263 purchasing observations in the toothbrush category of 775 households. In 27% of these observations the households previously purchased a toothpaste under umbrella A. All results were derived with the base prior specification (detailed in appendix A.1).

effect of choice inertia across categories would need to be far stronger than the effect within a category, and there is no reason to expect such a pattern.

The second and more important thing to note is that purchases in the toothpaste category actually have a significant influence on purchases in the toothbrush category: A look at the posterior distributions of the coefficients capturing cross-category inertia in umbrella brand choice (compare figure I.5) reveals that a previous purchase of a toothpaste under umbrella A increases the expected utilities of products under umbrella A and decreases those of products under other umbrella brands. The figures in table I.10 show that the effect of cross-category choice inertia is of similar size as the effect of within-category inertia. However, the number of observations suited for the identification of cross-category inertia is far smaller than the

γ^h: Inertia in umbrella choice	-0.58 (0.088)
β^h: Inertia in brand choice	-0.43 (0.089)
η^h: Price coefficient	-0.81 (0.043)
Effect of previous purchase of toothpaste under umbrella A on purchase of	
... toothbrushes under umbrella A (δ^h_A)	0.45 (0.148)
... toothbrushes under umbrella B or C ($\delta^h_{B,C}$)	-0.33 (0.106)
Marginal log-likelihood of...	
... model which accounts for cross-category inertia	-5989.285
... base model	-6064.658

Table I.10: **Summary statistics for posterior marginal distributions of model which accounts for cross-category inertia in umbrella brand choice.** The table gives summary statistics of the posterior marginal distributions depicted in figure I.5. The numbers in parentheses are the standard errors of the means. In the limit of infinite observations the numbers given would equal traditional frequentist maximum likelihood estimates of the model which accounts for cross-category inertia in umbrella brand choice (described in section I.7).

number of observations which can be used for the identification of within-category inertia. This is reflected in the relatively large standard errors of the means of the coefficients capturing cross-category inertia and advices caution when making quantitative comparisons.

In section I.6 I demonstrated that within-category inertia in umbrella brand choice can be explained by the existence of structural (or psychological) switching costs. In principle, I could repeat the analysis from section I.6 also for the dataset which contains information about toothpaste purchases. However, the necessary concentration on experienced households drastically reduces the number of observations available for the identification of cross-category inertia. Furthermore, inclusion of coefficients capturing cross-category inertia increases the degrees of freedom of my model. In effect, information from the data no longer dominates prior information, and statements regarding structural relationships can no longer be made. Thus, although it might seem suggestive that the same mechanisms are at play both for within- and across-category inertia, in order to give a definite answer to the question about the mechanism behind cross-category inertia in umbrella brand choice more data has to be collected.

I.8 Conclusion

In this chapter I analyzed how the widespread practice of marking products to be from one product family by the use of an umbrella brand influences households' purchasing behavior. I showed that households incur structural (or psychological) switching costs when they change umbrella brands. That is, besides the switching costs they incur from switching from one product to another, households incur additional switching costs when this other product is under another umbrella brand than the previously purchased product. These additional switching costs due to an umbrella brand change are of an economically significant size: I find that conditional on a product change a household ceteris paribus is indifferent between changing to a product also under the previously purchased umbrella brand and changing to a by 25% cheaper product outside the previously purchased umbrella brand. This effect of umbrella branding on households' purchasing behavior can be observed not only within a product category but also across product categories.

My findings add to the economic understanding of the practice of umbrella branding. So far, in the literature umbrella branding is primarily understood as a means of quality signaling (see for example Wernerfelt, 1988; Cabral, 2000, 2009; Hakenes and Peitz, 2008). The quality signal framework can rationalize household inertia in umbrella brand choice only as caused by learning about the quality of products linked by an umbrella brand. However, as shown in section I.6, also experienced households exhibit inertia in umbrella brand choice. In order to explain this fact one has to adapt the notion that apart from its quality signaling function umbrella branding also induces structural (respectively psychological) switching costs. One interpretation in the sense of Stigler and Becker (1977) would be that households choose among commodities which are composed of both the physical products and their brands - that is, the brand of a product is not only a signal of origin or affiliation to a product family, but an essential part of a household's consumption experience. This implies that households incur structural (or psychological) switching costs not only when they change physical products (as, amongst others, demonstrated by Dube et al., 2010), but also when they change umbrella brands.

The economic implications from the view of umbrella brands as quality signals are quite different to those from the view of umbrella brands as constituents of commodities which cause structural switching costs: First, in the quality signaling framework, umbrella branding only has dynamic implications as long as households are not fully informed about quality yet. In contrast, if umbrella brands are constituents of commodities which cause structural switching costs, the dynamic implications of umbrella branding are not restricted to a "learning period". Second, in the quality signaling framework the effects of umbrella branding are directly intertwined with product quality: In general, umbrella brands only assemble high-quality products. When umbrella brands are parts of commodities which cause switching costs, however, the effects of umbrella branding are not directly related to product quality. That is, whether households incur switching costs from changing umbrella brands is independent from the kind of products assembled under the umbrella brand.

Accordingly, whereas the existing theoretical literature on umbrella branding is mainly concerned with the interplay of a firm's decision whether to use umbrella branding and which quality to produce, theoretical studies adapting the view of umbrella brands as constituents of commodities which cause structural switching costs might focus more on the effects of umbrella branding on market structure. Klemperer (1995), for example, points out that the existence of switching costs might be a rationale for multi-product firms, as households who value variety and who incur firm-specific switching costs prefer multi-product firms over single-product firms. By analogy, the existence of switching costs due to the use of umbrella brands in the toothbrush category might explain the large variety of different toothbrush types assembled under each umbrella brand (up to 16, without counting different types of bristles and different sizes).

Chapter II

Information Disclosure in Non-Binding Auctions*

II.1 Introduction

When procuring a contract, the buyer often is not only interested in the price of an offer but also in other, non-price dimensions such as technical characteristics of the good or time of delivery. A by now quite well studied multidimensional auction format is given by scoring auctions where buyers prior to the bidding process establish a binding scoring rule. Besides such highly structured auctions, recently "non-binding" or "buyer-determined" auctions became increasingly important. In these auctions buyers can freely assign the contract after bidding has taken place. Currently this auction format seems to establish itself as the most prominent one for online marketplaces both for private and commercial contractors.[1]

When designing non-binding procurement auctions, typically no structure is imposed on the buyer's decision process - he is entirely free to choose any of the submitted bids. Important design questions arise, however, with respect to the optimal information structure for the bidding process. That is, bidders can be provided with different levels of information regarding the prices and the non-price characteristics of rival offers. Non-binding procurement auctions can be open-bid or sealed-bid auctions. If a non-binding auction is a sealed bid auction, bidders are usually neither informed about their rivals' prices nor their rivals non-price charac-

*This chapter is based on joint work with Gregor Zöttl.

[1]See Jap (2002, 2003), Jap and Haruvy (2008), and compare for example the platform FedBid, Inc., where US government agencies have procured more than $4.1 billion worth of purchases since 2008 using non-binding auctions.

teristics. However, in this chapter we do not deal with sealed-bid auctions. Instead, we are interested in open non-binding procurement auctions. In open non-binding auctions bidders are informed about their rivals' prices throughout the bidding process. The design question which arises here with regard to the information structure is whether information about their rivals' non-price characteristics is disclosed to or concealed from bidders.

In the present chapter we shed light on the optimal design of the information structure of open non-binding reverse auctions, using an extensive dataset from a large European online procurement platform. Our analysis focuses on the impact of transparency of the auction design with respect to bidders' non-price characteristics. In particular, we are interested in the effect of concealment of non-price information on the auction outcome. Theoretically, we find that the effect of concealment of non-price information depends on how the buyer weighs bidders' non-price characteristics against bidders' prices. We then do a counterfactual analysis to assess the relevance of this finding for applications in the field. If non-price information were concealed from the bidders, we would expect aggregate welfare of the buyers to decrease by up to 9% for auction-categories where buyers put only small weight on bidders non-price characteristics. In contrast, for auction-categories where buyers put a lot of weight on bidders' non-price characteristics we would expect aggregate welfare of the buyers to increase by up to 9%.

Our analysis proceeds as follows: First, we establish two different formal frameworks which describe two limiting cases of information structures. In the first case, bidders are fully informed about the non-price characteristics of their rivals. In the second case, all non-price information is concealed from the bidders. We show that whether it is beneficial for buyers to reveal non-price information depends on characteristics of the market considered, namely the relationship between the differences in the bidders' costs and that in their qualities, where a bidder's quality simply denotes how buyers value that bidder's non-price characteristics. The main intuition here is that when bidders are quite different in terms of how their non-price characteristics are valued by the buyers, then concealment of non-price information makes bidders appear more similar than they actually are, which toughens competition among bidders and in turn increases buyers' welfare.

Our empirical analysis is based on a detailed data set of an online procurement platform, where subscribed buyers post their tenders and can freely choose among the posted bids. For the period of observation all non-price information is publicly available to bidders. As a first step of our empirical analysis, for different auction categories we analyze how buyers value bidders' non-price characteristics. We then verify whether bidders indeed are aware of the buyers' preferences over their own and their rivals' non-price characteristics. Our theoretical frameworks imply that in this case, contrary to the case where non-price information is concealed from the bidders, the bids should directly take into account the non-price characteristics of rivals' bids. By exploiting the fact that a subset of bidders is observed to participate in several auctions we are able to identify the bidders' reactions to changing compositions of their rivals' non-price characteristics. We find that bidders submit significantly lower bids when confronted with rivals whose non-price characteristics are very valuable for the buyer.

After showing that bidders' observed behavior is indeed in line with our model for the case of disclosed non-price information, we conduct a counterfactual analysis and determine the impact concealment of quality information from the bidders would have on the welfare of the buyers for applications in the field. Using our model for the case of disclosed non-price information, we first derive estimates of the bidders' costs. We find that bidders' markups, which we compute using our cost estimates, are of expected size and in line with economic intuition - in particular, the average hourly profit is in the range of common net wages, and in auctions where bidders' qualifications matter markups are higher than in auctions in which jobs for low-skilled workers are procured. We then use these cost estimates together with our model for the case of concealed non-price information to compute bidders' counterfactual prices. With these we are finally able to calculate the change in the aggregate welfare of the buyers in case non-price information is concealed from the bidders. We do this for several job-categories which differ in the relevance of non-price characteristics. As we find, our theoretical predictions are of direct practical relevance for the dataset considered: For those job-categories where non-price characteristics are highly relevant (in our sample car repairs), buyers' welfare increases by up to 9%. In contrast, for those job-categories where non-price characteristics are of rather low importance (in our sample painting), buyers' welfare decreases by up to 9%.

Our work adds to a relatively new strand of literature which analyzes non-binding auctions. We are especially interested in the effect of different information structures in this auction format, however. There already are some interesting articles in this context. Several theoretical papers analyze the conditions under which it is beneficial for the buyer in non-binding auctions to inform bidders about their qualities. Gal-Or et al. (2007) show for sealed bid auctions that the buyer is better off when he discloses quality information to the bidders. Extensions such as the inclusion of risk averse bidders are provided in Doni and Menicucci (2010). Colucci et al. (2011) extend the setting of Gal-Or et al. (2007) by introducing heterogeneity in bidders' costs. They demonstrate that for the case of large cost differences and a comparatively small weighting of quality aspects it is in the best interest of the buyer to conceal quality information. In the opposite case, he is better off disclosing information about the bidders' quality.[2]

In a recent article, Haruvy and Katok (2013) are the first ones to shed more light on those issues from an empirical perspective. Based on controlled laboratory experiments, they analyze both open and sealed bid non-binding auctions and assess the impact of information revelation on bids submitted. For the parameter environments chosen in their laboratory experiments they find that in their open auction design due to more aggressive bidding buyers are better off if they keep information about bidders' qualities concealed. Our work differs from their contribution since our analysis is based on field data of indeed conducted auctions. Our analysis, moreover, is conducted for several different services to be procured, and thus allows us to identify under which conditions information revelation indeed is desirable in open non-binding auctions. That is, for the case of car repairs our results are in line with those obtained by Haruvy and Katok (2013), whereas for the case of painting we obtain opposite results.

Several recent articles compare the performance of non-binding auctions with regular price only auctions. Engelbrecht-Wiggans et al. (2007) is one of the seminal articles in this context. They analyze both analytically and experimentally under which conditions the buyer would want to commit to a price only mechanism which ignores all non-price attributes. As the authors establish, such commitment is only desirable when competitive pressure is important (few bidders) and expected quality of the low-cost-bidders is not too low (limited negative correlation between cost and quality).[3] Fugger et al. (2013) find in a recent contribution that when bidders

[2]For a similar setting Rezende (2009) shows that when the buyer and the suppliers have the possibility to renegotiate, it can be optimal for the buyer to fully reveal the information about the suppliers' qualities.

[3]In principle, also our setting compares a non-binding auction (with informed bidders) with a "price-only"-regime. In our setting, however, "price-only" refers solely to the information held by the bidders, who know that prices matter, but are uncertain with respect to all other criteria. The buyers always do choose the ex post best offer, taking into account all non-price characteristics (as in Gal-Or et al., 2007; Doni and Menicucci, 2010; Haruvy and Katok, 2013). The fundamental insights obtained in our analysis are thus clearly quite different. As one consequence, for example,

are uncertain about the exact way different criteria enter the final decision of the buyer, there are cases where a non-binding auction enables them to coordinate on high prices. In that case the buyer would prefer binding price-only auctions over non-binding auctions.

Wan and Beil (2012) and Wan et al. (2012) analyze related but slightly different problems. They study auctions where bidders in order to win the auction additionally have to meet certain quality standards. Those articles explore theoretically and experimentally under which conditions it is optimal to provide information with respect to the screening among bidders either prior or after bidding has taken place.

Our work in general contributes to the literature which analyzes efficient ways to procure contracts when the buyer's valuation of an offer depends on additional dimensions besides price. Scoring auctions (where binding scoring rules take price and non-price characteristics into account) have already received significant attention in the literature and by now are quite well understood. Asker and Cantillon (2008, 2010) show that for the case when suppliers have multi-dimensional private information, this procurement mechanism dominates others like sequential bargaining and price-only auctions. Different scoring auction designs are compared in Che (1993), Branco (1997), Chen-Ritzo et al. (2005) and Kostamis et al. (2009). Empirical analysis of scoring auctions can be found in Athey and Levin (2001) and Lewis and Bajari (2011), the first using data from US timber auctions and the second data from US highway procurement auctions. Practical implementability of scoring auctions through iterative process is analyzed for example in Bichler and Kalagnanam (2005) or Parkes and Kalagnanam (2005).Finally, in a theoretical contribution Che (1993) compares scoring auctions with non-binding auctions. He shows that when bidders bid on all dimensions of their offers, from the buyer's perspective scoring auctions dominate non-binding auctions.

The chapter proceeds as follows. Section II.2 introduces our theoretical frameworks for the case of disclosed and that of concealed non-price information and derives under what conditions a buyer prefers which information regime. Section II.3 introduces our dataset. In section II.4, for different auction categories we analyze how buyers value bidders' non-price characteristics, and in section II.5 we use a reduced-form model to show that bidders' behavior is indeed in line with our framework for the case of disclosed non-price information. Based on these preparations, in section

the correlation between cost and quality, which is crucial in Engelbrecht-Wiggans et al. (2007), is not of central importance in our setting since foregone quality is not an issue.

II.6 we perform a counterfactual analysis to assess how strongly buyers' welfare can be expected to change if non-price information gets concealed from the bidders. This is done for several auction categories. Section II.7 concludes.

II.2 Theoretical Framework

■ **Framework.** We consider a non-binding and open procurement situation where a buyer wants to procure some contract among J participating firms. Each firm has some cost c_j for providing the service. For the sake of exposition, we initially assume that costs are known among firms. Below we will show that all our results also hold in case costs are unknown but bidders are myopic or as one possible equilibrium in case costs are unknown and bidders are perfectly rational. Bidding takes place throughout different rounds $r = 1, ..., R$ of the auction. In each round, each firm $j = 1, ..., J$ can successively update its publicly observable price-bid $b_{j,r}$. We denote the vector of final bids $b_{j,R}$ quoted by each firm once bidding has stopped by $\mathbf{p} = (p_1, ..., p_J)$. Once price submission has finished the buyer can freely choose to award the contract to some firm j at price p_j.

For the buyer's decision not only the final price p_j quoted by firm j matters but also its non-price characteristics, which we denote by $\mathbf{A}_j$ and which we assume to be exogenously given. In analogy to the existing literature on non-binding procurement auctions, we call the value of these non-price characteristics to the buyer a firm's *quality* q_j. Given the buyer's preferences regarding these non-price characteristics, which we denote by $\boldsymbol{\alpha}$, we assume that the quality of firm j is a linear function of that firm's non-price characteristics, $\mathbf{A}_j$, and the respective preferences $\boldsymbol{\alpha}$ of the buyer. That is, $q_j = \boldsymbol{\alpha}\mathbf{A}_j$.

Throughout our analysis, we assume that the buyer is always fully informed about each firm's non-price characteristics. However, with respect to the *information firms receive about other firms' non-price characteristics* we differentiate between two cases: In the first case, which we call *information case (IC)*, each firm is informed about each other firm's non-price characteristics. That is, in the information case $\mathbf{A} = (\mathbf{A}_1, ..., \mathbf{A}_J)$ is common knowledge. In the other case, which we call *no information case (NIC)*, the firms are not informed about each other's non-price characteristics. That is, in the no information case firm j does not know about the non-price characteristics $\mathbf{A}_k$ of each of his rival firms k.

We assume that the buyer can choose among J firms and an outside option. He receives a certain amount of utility u_j when he chooses firm j. This amount of utility depends on the price p_j put forward by this firm and the firm's exogenous non-price characteristics $\mathbf{A}_j$. We model the utility a buyer receives from a certain firm as being linearly dependent on the price p_j, the firm's non-price characteristics $\mathbf{A}_j$, and an error term ϵ_j. With that, we assume the buyer's decision process to be given as

$$\begin{aligned} \max_{j\in\{0,1,\ldots,J\}} \quad & u_j, \quad \text{where} \\ u_0 &= v_0 + \epsilon_0 \\ u_1 &= -p_1 + \boldsymbol{\alpha}\mathbf{A}_1 + \epsilon_1 \\ &\vdots \\ u_J &= -p_J + \boldsymbol{\alpha}\mathbf{A}_J + \epsilon_J \end{aligned} \tag{II.1}$$

$\boldsymbol{\alpha}$ denotes the vector of the buyer's preferences regarding firms' non-price characteristics. v_0 denotes the value of the buyer's outside option. For simplicity and without loss of generality we normalize the price coefficient to -1. The error terms ϵ_j capture uncertainty in the buyer's decision due to unobserved influences unrelated to price or non-price characteristics.[4] We assume that the ϵ_j follow a symmetric distribution with mean zero. When making his decision, the realizations of the ϵ_j are known to the buyer, but they always remain concealed from the firms while bidding. The buyer is assumed to choose the option which maximizes his utility, that is, the option k for which

$$u_k > u_j \qquad \forall j \neq k, \qquad j,k \in \{1,\ldots,J\}.$$

□ **Information case.** We assume that in the information case firms have full information about both their own and their rivals' non-price characteristics $\mathbf{A}$. In a non-binding auction, in contrast to a scoring auction, there is no binding and publicly known scoring rule. That is, firms are not explicitly informed about the way the buyer makes his decision. Instead, we assume that firms collected information about the buyer's decision process by observing past auctions. Thus, each firm's

[4] For example, the buyer might be influenced in his decision by his (unobserved) taste regarding, for example, the username firm j chooses at a bidding platform.

model of the buyer's decision process is given as

$$\begin{aligned} \max_{j \in \{0,1,\ldots,J\}} \quad & u_j, \quad \text{where} \\ u_0 &= v_0 + \epsilon_0 \\ u_1 &= -p_1 + \boldsymbol{\alpha}\mathbf{A}_1 + \epsilon_1 \\ &\vdots \\ u_J &= -p_J + \boldsymbol{\alpha}\mathbf{A}_J + \epsilon_J. \end{aligned} \tag{II.2}$$

Note that in contrast to the buyer, who knows the realizations of the ϵ_j when making his decision, from the firms' perspectives the ϵ_j are random. We assume that the unobservables ϵ_j follow some distribution, and that the firms know the distribution of the ϵ_j. In consequence, given some bid p_j of its own, firm j can derive all winning probabilities $P_k(\mathbf{p}, \mathbf{A})$, $k \in \{0, 1, ..., J\}$. These winning probabilities are functions of all firms' final price bids $\mathbf{p} = (p_1, ..., p_J)$ and all firms' non-price characteristics $\mathbf{A} = (\mathbf{A}_1, ..., \mathbf{A}_J)$. We assume that the winning probability P_k of each firm k is log concave in its final price quote.[5] Expected profits π_j of firm j are given by

$$\pi_j = P_j(\mathbf{p}, \mathbf{A})(p_j - c_j).$$

Within our framework, we obtain a unique subgame perfect Nash equilibrium $\mathbf{b}^*$ of submitted bids. Since throughout the chapter only final price bids are relevant for the decison of the buyer, we explicitly focus on the final price bids $\mathbf{p}^*$ arising within this equilibrium. These are given as the mutually best responses to the final price bids of all rivals and are characterized by the following expression:

$$p_j + \frac{P_j}{\partial P_j / \partial p_j} - c_j = 0, \quad \forall j \in \{1, ..., J\}. \tag{II.3}$$

The winning probabilities P_j follow from (II.2) and depend on all bidders' prices $\mathbf{p}$ and non-price characteristics $\mathbf{A}$. Existence and uniqueness of $\mathbf{p}^*$ as characterized by (II.3) has already been shown in the literature, compare Caplin and Nalebuff (1991) and Mizuno (2003).

□ **No information case.** We assume that in the no information case firms are not informed about each other's non-price characteristics $\mathbf{A}_j$. Analogous to the information case we assume that there is no binding and publicly known scoring rule, but that firms instead had to collect information about the buyer's decision process from observing past auctions. As in the no information case non-price

[5] Notice that the logit framework referred to from section II.6 onwards satisfies this assumption.

information is concealed, in their model of the buyer's decision process firms can only take into account the observable prices. That is, we assume firms' model of the buyer's decision process to be:

$$\begin{aligned}
\max_{j\in\{0,1,\dots,J\}} \quad & u_j, \quad \text{where} \\
u_0 &= \tilde{v}_0 + \tilde{\epsilon}_0 \\
u_1 &= -p_1 + \tilde{\epsilon}_1 \\
&\vdots \\
u_J &= -p_J + \tilde{\epsilon}_J.
\end{aligned} \tag{II.4}$$

Note that in contrast to the buyer firms do not know about the realizations of the terms $\tilde{\epsilon}_j$. Given final price bids $\mathbf{p}$, firm j can derive winning probabilities $\tilde{P}_k$, $k \in \{0, 1, ..., J\}$. These winning probabilities are functions of only the firms' price bids. We assume that the winning probability $\tilde{P}_k$ of each firm k is log concave in its final price quote.[6] The expected profit $\tilde{\pi}_j$ of firm j is given by

$$\tilde{\pi}_j = \tilde{P}_j(\mathbf{p}) \cdot (p_j - c_j)\,. \tag{II.5}$$

In analogy to the information case discussed above we obtain a unique subgame perfect Nash equilibrium $\mathbf{b}^*$ also for the no information case. The final price bids $\mathbf{p}^*$ arising within this equilibrium are mutually best responses to the final price bids of all rivals. They are characterized by the following expression:

$$p_j + \frac{\tilde{P}_j}{\partial \tilde{P}_j / \partial p_j} - c_j = 0, \quad \forall j \in \{1, ..., J\}. \tag{II.6}$$

That is, the equilibrium is obtained analogous to the information case. However, the winning probabilities as perceived by the bidders, $\tilde{P}_j$, are now determined by (II.4) and depend only on bidders' final price bids $\mathbf{p}$.

■ **Robustness - alternative framework formulation.** When establishing tractable and thus necessarily stylized frameworks it is always highly debatable whether the chosen framework best approximates the real world economic interaction in a meaningful way. In the context of modelling auctions, this problem clearly is much more pronounced when analyzing open and thus in principle dynamic auction settings (the case of the present chapter). Let us emphasize that the final price quotes (as characterized in expressions (II.3) and (II.6)) arise as an equilibrium outcome also

[6]Notice that the logit framework referred to from section II.6 onwards satisfies this assumption.

under modified assumptions for example with respect to the precise endpoint of the auction and with respect to the information structure for procurement costs.

To this end consider for example the framework where each firm has some private costs c_j with distribution $F_j(c)$ with positive and bounded support for providing the service. Bidders $j = 1, ..., J$ can successively update their publicly observable price-bids $b_{j,r}$ throughout different rounds $r > 2$ of the auction. The bidding stops when neither of the bidders updates his price bid in a specific round.

Notice that the final price vector p^* characterized in expressions (II.3) and (II.6) respectively results as the equilibrium outcome of a perfect Bayesian equilibrium of the induced auction game. An equilibrium strategy profile supporting this outcome is given as follows: Whenever it is a bidder's turn he sets his currently active bid such as to be a best response to the currently active bids of all rival bidders (clearly this is also well defined in case no or only few active bids are already submitted when bidding just started). Beliefs with respect to rivals' costs are updated consistently given the observed bids of rivals. The above characterized final price vector results from those equilibrium strategies when none of the bidders wants to update his bid any more. The fact that for given p^* none of the bidders wants to update his bid any more obtains by construction, since each price is the best response given all rivals' price bids. Notice, however, that for the case of fully rational bidders also other equilibria of the above specified auction framework can obtain. Consider, for example, some vector of (collusive) prices p^C which for each bidder (and each cost type) grants expected profits strictly above those obtaining for p^*. A perfect Bayesian strategy profile supporting the above characterized equilibrium outcome p^C is given as follows: Independently of their cost type bidders in the first round submit prices p_j^C, in the second round no bids are submitted and the auction stops. If any bidder deviates from this strategy, all bidders will start to submit bids such as to be a best response to all currently active bids.

Finally notice, for the case of myopic bidders (and also for fully rational bidders which consider all their rivals to behave myopically[7]) we obtain the final price vector p^* characterized in expressions (II.3) and (II.6) as the unique outcome of the above specified alternative framework. As shown by Sobel and Wei (2010) equivalent results obtain when restricting attention to markov perfect equilibria of the dynamic auction game.

[7] Empirical evidence that players indeed significantly underestimate their rivals' rationality can be found, for example, in Weizsaecker (2003).

In summary, we see that also for other reasonable framework and equilibrium refinement choices the price vector p^* characterized in expressions (II.3) and (II.6) always arises in equilibrium. Not surprisingly, for deliberately general settings of dynamic and strategic interaction among bidders uniqueness of the equilibrium outcome cannot always be guaranteed for all in principle plausible settings, however.

■ **Comparing both information regimes.** We are interested in whether the buyer prefers to disclose or to conceal non-price information. We assume that this decision has to be made prior to knowing the precise number and identity of the participating firms and their characteristics. In this case, the buyer prefers the information structure which gives him the highest expected utility. It is easy to show that *there is no information structure which dominates the other.* Appendix B.1 proofs this by example.

The central intuition is that the informational arrangement which creates the highest competitive pressure among firms is best for the buyer. Which information regime creates more competitive pressure as perceived by the firms depends on the specific situation considered, as we show. First, consider a situation where firms have similar production costs but are quite heterogenous with respect to how the buyer values their non-price characteristics. In short, using the definition of a firm's quality as the buyer's valuation of its non-price characteristics, that means a situation where firms have similar production costs but very different qualities. A regime which conceals non-price information suggests tough competition and induces more aggressive bidding. Second, consider a situation where firms have quite different production costs but quality differences are such as to compensate for those differences (that is, the more expensive producer also has higher quality). In this case full revelation of non-price information induces more aggressive bidding. In the following section we offer an analytical illustration of these tradeoffs.

■ **Illustration of tradeoffs and model mechanics.** The standard assumption about the distribution of the ϵ_j is either normal or type I extreme value. Bidders' winning probabilities P_j depend on the differences of the ϵ_j. In case the error terms are assumed to follow a normal distribution also their differences follow a normal distribution, and in case the error terms are assumed to follow a type I extreme value distribution their differences follow a logit distribution. In consequence, the P_j either cannot be expressed in closed form or contain exponential terms which lead to transcendental equations. Thus, for any standard assumption about the distribution of the ϵ_j the first order conditions (II.3) respectively (II.6) cannot be solved analytically.

To illustrate the mechanics of our model we make the simplifying assumption that the differences of the error terms ϵ_j follow a uniform distribution. With this assumption we analyze bidding in an auction where the buyer can choose among two firms only.[8] Each firm has non-price characteristics $\mathbf{A}_j$. The respective preferences of the buyer are denoted by $\boldsymbol{\alpha}$. The buyer's valuation of a firm's non-price characteristics, that is its quality q_j, is given as $q_j = \boldsymbol{\alpha}\mathbf{A}_j$. We assume that firm 1 is of low quality and low costs, while firm 2 is of high quality and high costs. That is, $q_1 < q_2$ and $c_1 < c_2$. $\epsilon_2 - \epsilon_1$ shall follow a uniform distribution with mean zero and variance ν. $\tilde{\epsilon}_2 - \tilde{\epsilon}_1$ shall follow a uniform distribution with mean zero and variance $\tilde{\nu}$. As in the no information case bidders are missing non-price information they should perceive the buyer's decision to be more noisy. Thus, we assume that $\tilde{\nu} \geq \nu$. With these assumptions it is possible to derive illustrative analytical results. Their derivation can be found in appendix B.2.

☐ **Relationship between firms' equilibrium bids.** By making use of the first order conditions (II.3) and (II.6) we derive equilibrium bids for the information case and the no information case. We directly turn towards the comparison of the equilibrium bids. The p_i denote the equilibrium bids in the information case, the $\tilde{p}_i$ the equilibrium bids in the no information case:

$$p_1^* = \tilde{p}_1^* - \frac{1}{3}(q_2 - q_1) - \sqrt{3}(\sqrt{\tilde{\nu}} - \sqrt{\nu}), \tag{II.7}$$

$$p_2^* = \tilde{p}_2^* + \frac{1}{3}(q_2 - q_1) - \sqrt{3}(\sqrt{\tilde{\nu}} - \sqrt{\nu}). \tag{II.8}$$

The intuition behind expressions (II.7) and (II.8) is straightforward: The first term added to $\tilde{p}_2^*$ respectively subtracted from $\tilde{p}_1^*$ captures that in case of disclosed non-price information firms become aware of firm two's competitive advantage in terms of quality: The net competitive pressure on the low-quality firm (firm one) increases, while that on the high-quality firm (firm two) decreases. The last term in expressions (II.7) and (II.8) captures that in case of concealed non-price information firms perceive the buyer's decision to be more noisy and thus add a markup on their costs.

☐ **Relationship between buyer's expected utilities.** The relationship between the expected utility of the buyer in the information case, EU, and that in the no

[8] We implicitly assume that the value of the outside option is so low that the upper limit to the prices of firm 1 and 2 is above the equilibrium prices. The outside option simply leads to upper limits for the prices of firm 1 and 2. Thus, its explicit treatment would only make our analysis more complicated without delivering further insights.

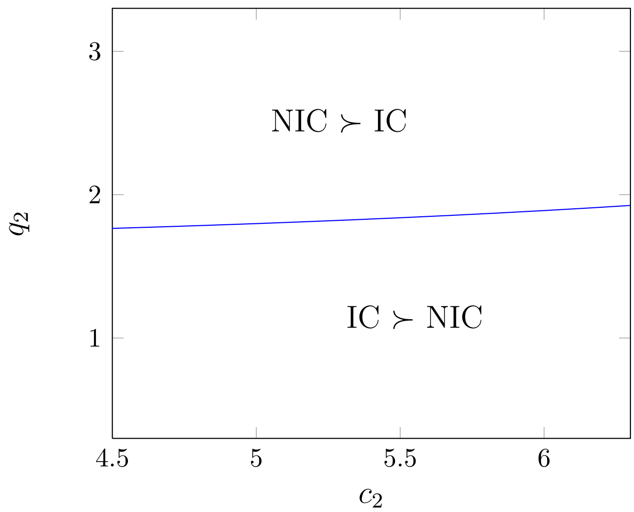

Figure II.1: **Preferences of the buyer regarding the information structure as a function of the auction parameters.** The graph shows the indifference line of the buyer. The indifference line of the buyer represents the parameter set at which the buyer is indifferent between disclosing and concealing non-price information. We assumed that $c_1 = 4.5$ and $q_1 = 0.3$. The chosen parameters sizes resemble typical parameter sizes from our application. For all q_2-c_2-combinations above the indifference line the buyer prefers to conceal non-price information, whereas for all combinations below he prefers to disclose non-price information.

information case, $\widetilde{\mathrm{EU}}$, is given as

$$\begin{aligned} \mathrm{EU} - \widetilde{\mathrm{EU}} \quad = \quad & \frac{1}{3\sqrt{12\nu}}(q_2 - q_1)\left[(c_2 - c_1) - 2(q_2 - q_1)\right] \\ & +3(2\sqrt{\nu\tilde{\nu}} + \tilde{\nu} - 3\nu) \\ & +(\frac{\sqrt{\tilde{\nu}}}{2\sqrt{\nu}} - \frac{1}{2})(c_2 + c_1 - q_2 - q_1). \end{aligned} \tag{II.9}$$

Equation (II.9) shows that the net change in the expected utility of the buyer depends on three factors: The first term captures the tradeoff between the competitive advantage of the low-cost firm and that of the high-quality firm. If the difference in costs is small but that in qualities is very high, disclosure of non-price information weakens competition because firms become aware of the high-quality firm's large net advantage. If in contrast the difference in costs is very high and that in qualities small, disclosure of non-price information strengthens competition as it mitigates the net advantage of the low-cost firm. The second term captures that in the no information case firms perceive the decision of the buyer to be more noisy. In the no information case they thus demand a markup on their prices which in turn decreases buyer's welfare. The third term weighs the effect of uncertainty (term two) against that of quality information (term one). The weight of either effect depends on how strong relative to costs firms' pricing decisions are influenced by non-price information. The smaller the influence of non-price information, the more the effect of noise in the buyer's decision outweighs that of non-price information.

The graph in figure II.1 illustrates how the buyer's preferences regarding the information structure change as a function of the auction parameters, namely the firms' costs and their qualities. The parameter sizes used for this example resemble typical parameter sizes from our application.[9] The important take-away is that which information structure to choose for a certain application is not clear ex ante but depends on the setting. In general, if the difference in qualities is high and that in costs is low, the buyer prefers the no information case over the information case. In contrast, if the difference in qualities is low and that in costs is high, the buyer prefers the information case over the no information case.

II.3 Data

We have available an extensive dataset from a popular European online procurement platform. On this platform private customers tender jobs ranging from construction over general repair and renovation to teaching. Jobs are awarded through an open non-binding auction.

The exact procedure is as follows: A private customer (the buyer) posts a description of the job he wants to procure. This description is entered into a free-text field and usually contains job details (for example, the area to be painted, whether or not cleaning is required, and so on), the job site, a price expectation (termed "startprice" in the following) and an announcement of the time span during which tradesmen (the bidders) can put forward quotes. All this information is available to all tradesmen registered at the platform. During the defined time span all interested tradesmen can publicly announce prices for which they are willing to do the offered job. Announced prices can be changed at any point during the auction. The current price quote of each bidder and all his non-price characteristics are publicly observable on the website. The non-price characteristics of a bidder include the number of positive and negative ratings the bidder received so far, his home location, qualifications like the possession of certain degrees, his area of expertise, and so on. At the end of the auction the buyer is free to award the job to one of the bidders or to withdraw his

[9] The average value of $\boldsymbol{\alpha}\mathbf{A}_j$ in our data (that is, the average quality q_j) is 0.3. The average (estimated) costs are €450. In "utility-units" this is 4.5 (which equals the average value of ρc_j). For our example, we set $c_1 = 4.5$ and $q_1 = 0.3$.

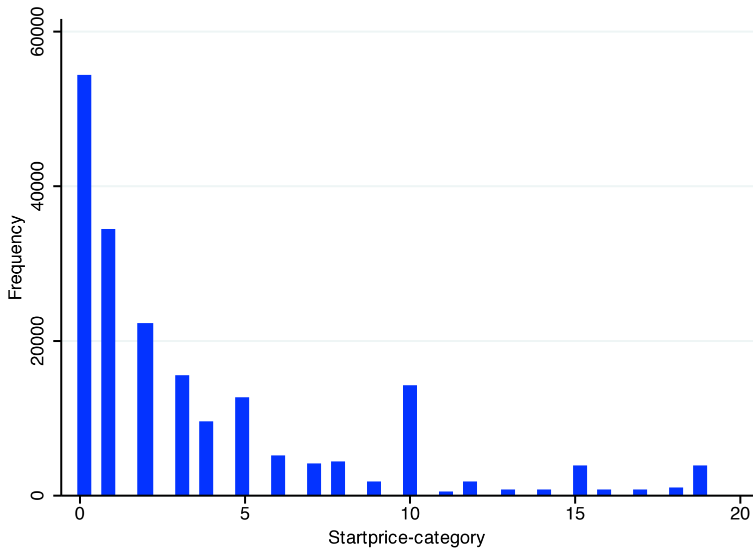

Figure II.2: **Distribution of auctions over startprice categories.** Displayed is the distribution of all auctions which were conducted during the second half of 2008 (187,747 auctions) over startprice-categories. Startprice-category 0 ranges from €0-99, startprice-category 1 ranges from €100-199, and so on.

offer. In case of an award the platform obtains a certain percentage of the successful bid as commission.

We have available data on auctions which were conducted during the years 2007 and 2008. In this time span the auction platform experimented with some rule changes. In order to exclude the possibility that our results are influenced by these rule changes we focus our analysis on auctions which took place during the second half of the year 2008. In this period there were only minor rule changes. like for example a slight reduction of the time span after which the buyer has to decide whether to withdraw his offer or award the job to one of the participating bidders. Minor changes like these should have no effect on our results.

In the second half of 2008 we observe around 180,000 auctions. These are grouped into a number of categories like “painting”, “moving”, and so on. Besides by the kind of job offered auctions are differentiated by the value of the jobs offered. We use the price expectation the buyer states at the beginning of the auction (the startprice) as a proxy for the value of the job offered.[10] Startprices can be chosen freely but

[10]The level of the startprices put forward by the buyers is highly correlated with the level of the prices the bidders put forward, which reassures us that startprices are indeed good proxies for the value of the jobs procured. Note also that the startprice is set purely for informational reasons, it neither puts any restriction on bids submitted nor on the award decision made by the potential buyer.

are typically set in multiples of €100 (respectively €50 for auctions with values below €500). More than half of all auctions have a startprice which is below €300. We sort the auctions into different startprice-categories: Category 0 ranges from €0-99, category 1 from €100-199, and so on. Figure II.2 depicts the distribution of the auctions over these startprice-categories. As can be seen, nearly 30% of all auctions have startprices between €0-99, and of all these auctions with startprices between €0-99 around 80% have a startprice of €50 or less. We expect the bidding behavior in these very low valued auctions to be fundamentally different from the bidding behavior in auctions with higher stakes and thus drop all auctions from startprice-category 0 from our analysis.

For every auction in each job-startprice-category we have available information about the number and the identities of the participating bidders, the prices put forward, the bidders' non-price characteristics (like the number of positive and negative ratings, the possession of certain degrees and qualifications, and so on) and the final choice of the potential buyer (including whether he chose to withdraw his job offer). For our analysis we focus on the four top categories with respect to the number of auctions. These are "moving", "painting", "car" and "plumbing and heating". We use only auctions in which at least two bidders participate. Descriptive statistics for each auction-category are given in table II.1. The left part of figure II.3 shows the spatial distribution of all auctions conducted, the right part gives an exemplary impression of the course of an auction.

As already mentioned, on the auction platform we have our data from both buyers and bidders are fully informed about each bidders' non-price characteristics. We are interested in what would happen to the welfare of the buyers if these non-price information were concealed from the bidders. Our theoretical considerations in section II.2 show that, among other things, the answer depends on how buyers value bidders' non-price characteristics. We think it is reasonable to expect buyers' preferences $\boldsymbol{\alpha}$ regarding bidders' non-price characteristics to depend both on the job category and on the value of the job offered. For example, whether a bidder has undergone professional training should matter more for jobs from the "plumbing and heating" category than for jobs from the "moving" category. Similarly, whether a bidder has liability insurance might matter more for a buyer when he procures a high-value job than when he procures a low-value job. To capture that the choice behavior of a buyer (and in consequence the behavior of the bidders) possibly depends on the type and the value of the job offered, we will perform separate analyses for each of the four most frequent job categories ("moving", "painting", "car", "plumbing and

"Moving"	Mean	SD	Median	Min	Max
Nbr. of auctions	16,841				
Nbr. of bidders	4,564				
Nbr. of buyers	15,076				
Nbr. of bidders per auction	5.1	3.1	4	2	27
Bid amount	556.7	463.7	450	1	3000
Nbr. of auction participations per bidder	5.3	35.8	1	1	1748
Auctions per buyer	1.1	0.5	1	1	23
Auction duration (days)	10.6	9.6	8.7	0	144.0
Last bid placement (hours till auction end)	88.5	160.8	20.0	0	1,883.7

"Painting"	Mean	SD	Median	Min	Max
Nbr. of auctions	11,434				
Nbr. of bidders	5,800				
Nbr. of buyers	10,614				
Nbr. of bidders per auction	6.4	4.2	5	2	31
Bid amount	606.2	496.3	450	0	3000
Nbr. of auction participations per bidder	5.2	21.9	1	1	793
Auctions per buyer	1.1	0.3	1	1	8
Auction duration (days)	11.4	9.3	10	0	120.0
Last bid placement (hours till auction end)	84.2	162.4	12.3	0	1,891.8

"Car"	Mean	SD	Median	Min	Max
Nbr. of auctions	3,413				
Nbr. of bidders	1,541				
Nbr. of buyers	3,216				
Nbr. of bidders per auction	2.8	1.2	2	2	12
Bid amount	398.0	449.4	250	1	3000
Nbr. of auction participations per bidder	2.7	12.4	1	1	397
Auctions per buyer	1.1	0.3	1	1	4
Auction duration (days)	15.3	12.1	14	0	118.1
Last bid placement (hours till auction end)	150.8	215.5	53.2	0	1,786.7

"Plumbing and Heating"	Mean	SD	Median	Min	Max
Nbr. of auctions	2,962				
Nbr. of bidders	2,161				
Nbr. of buyers	2,831				
Nbr. of bidders per auction	2.9	1.3	2	2	13
Bid amount	471.4	604.9	198	1	3000
Nbr. of auction participations per bidder	2.3	6.5	1	1	156
Auctions per buyer	1.0	0.3	1	1	7
Auction duration (days)	13.1	10.9	10.0	0.1	109.4
Last bid placement (hours till auction end)	112.2	183.1	30.3	0	1,518.3

Table II.1: **Descriptive statistics for auctions from job categories "moving", "painting", "car" and "plumbing and heating".** The table displays descriptive statistics for auctions from the four most popular job categories ("moving", "painting", "car" and "plumbing and heating"). Considered are all auctions with startprices ranging from €0-2000 and with at least two participating bidders.

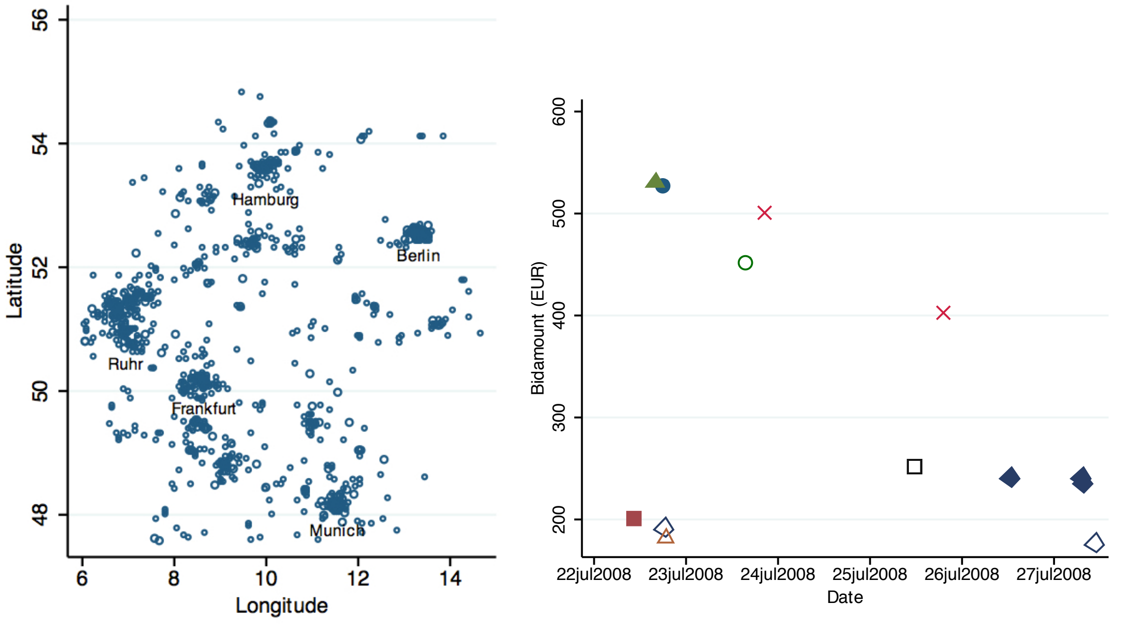

Figure II.3: **Spatial distribution of auctions and exemplary bidding process.** On the left, the spatial distribution of all auctions from job category “painting” for which we have collected cost information is displayed. On the right, an example of a typical bidding process is shown. The different symbols stand for different bidders, the auction is from the job category “painting”, and the startprice set was €200.

heating”) and for each of the four most frequent startprice-categories (1, 2, 3 and 5).

For a sample of around two-thousand auctions from the job category “painting” we manually extracted information about cost factors from the job descriptions. These cost factors include for example the area to be painted, whether paint is provided by the buyer, and so on. We do not need this information for our counterfactual analysis in section II.6, where for each job-startprice-category we analyze the change in aggregate welfare of the buyers in case non-price information gets concealed. However, before doing our counterfactual analysis in section II.6, in section II.5 we use information about cost factors to verify a fundamental assumption of our counterfactual analysis. This assumption is that bidders know about the preferences of the buyers regarding their non-price characteristics and that thus their behavior is in line with our information case model.

II.4 Analysis of Buyers' Preferences

Besides price information buyers have available information about the non-price characteristics of the bidders. We assume that when making their decisions buyers use both price and non-price information. In particular, we assume that a buyer's ranking of a given bidder depends on both the price that bidder puts forward and how he values that bidder's non-price characteristics. With $\mathbf{A}_j$ denoting the vector of bidder j's non-price characteristics and $\boldsymbol{\alpha}$ denoting the vector of the buyer's respective preferences, we assume the buyer's valuation of bidder j's non-price characteristics (that is, bidder j's quality) is equal to $\boldsymbol{\alpha}\mathbf{A}_j$. We observe each bidder's non-price characteristics, but we do not observe the preferences of the buyers. In this section we use a logit discrete choice model to elicit buyers' preferences $\boldsymbol{\alpha}$. We do this in order to see whether buyers indeed behave in accordance with our assumption that they account for non-price information when making their decisions. Furthermore, information about buyers' preferences is at the core of our counterfactual analysis in section II.6, as we assume the bidders to base their behavior on a rational model of buyers' behavior.

■ **Econometric model.** For a given auction n we model the decision of the buyer as a discrete choice among all participating bidders and an outside option. We assume the buyer to base his decision among bidders on both their prices and their non-price characteristics. Bidders' non-price characteristics comprise binary characteristics, indicating for example the possession of certain degrees, discrete characteristics like the number of positive and negative ratings, and a continuous measure for the distance between a bidder's home location and the job site.[11]

We estimate buyers' preferences along the lines of the model we developed in section II.2: In a given auction n, buyer's utility from choosing bidder j is assumed to be linearly dependent on the bidder's price p_{nj}, how he values the bidder's non-price characteristics, and an error term ϵ_{nj}. We assume that the buyer's valuation of a bidder's non-price characteristics is a linear function of that bidder's non-price characteristics and the buyer's respective preferences. With $\mathbf{A}_{nj}$ subsuming the non-price characteristics of bidder j in auction n, and $\boldsymbol{\alpha}$ denoting the preferences of the buyer over these characteristics, the buyer's valuation of the non-price characteristics

[11] The distance measure is constructed from the buyers' and the bidders' zip-codes. As such it is only approximate. However, given the assumption that also the buyers can in general be expected to base their decision on a rough distance estimate and not an exact calculation, it should suffice to capture the respective part of the buyers' decisions.

of bidder j in auction n is given by $\boldsymbol{\alpha}\mathbf{A}_{nj}$.[12] With ρ denoting the price elasticity of the buyer in auction n,[13] the utility he derives from each of the J_n participating bidders can explicitly be formulated as

$$\begin{aligned} u_{n0} &= \epsilon_{n0} \\ u_{n1} &= t + \rho p_{n1} + \boldsymbol{\alpha}\mathbf{A}_{n1} + \epsilon_{n1} \\ &\vdots \\ u_{nJ_n} &= t + \rho p_{nJ_n} + \boldsymbol{\alpha}\mathbf{A}_{nJ_n} + \epsilon_{nJ_n}. \end{aligned} \tag{II.10}$$

The constant t captures the value of the outside option. It holds that the lower t the higher is the value of the outside option. The error terms ϵ_{nj} capture unobserved influences on the buyer's decision unrelated to bidders' prices or their observed non-price characteristics. The buyer is assumed to choose the bidder which offers him the highest utility. By assuming the error terms ϵ_{nj} to be independently, identically type I extreme value distributed we obtain the standard logit model: The choice probabilities are given as

$$P_{nj} = \begin{cases} \frac{1}{1+\sum_{k=1}^{J_n} e^{t+\rho p_{nk}+\boldsymbol{\alpha}\mathbf{A}_{nk}}} & \text{if } j = 0, \\ \frac{e^{t+\rho p_{nj}+\boldsymbol{\alpha}\mathbf{A}_{nj}}}{1+\sum_{k=1}^{J_n} e^{t+\rho p_{nk}+\boldsymbol{\alpha}\mathbf{A}_{nk}}} & \text{if } j \in \{1, ..., J_n\}. \end{cases}$$

Estimates of the model parameters $\{\rho, \boldsymbol{\alpha}\}$ can be obtained by maximizing the likelihood

$$L = \prod_{n=1}^{N} \prod_{j=0}^{J_n} (P_{nj})^{y_{nj}}, \quad y_{nj} = \begin{cases} 1 & \text{if alternative } j \text{ is chosen in auction } n, \\ 0 & \text{otherwise.} \end{cases}$$

■ **Estimation results.** We estimate our model for each combination of the job categories “moving”, “painting”, “car”, “plumbing” and the startprice-categories 1, 2, 3, 5. Table II.2 displays the results for startprice category 1 (which covers all auctions with startprices ranging from €100-199) and all job categories. Table II.3 displays the results for job category “moving” and all startprice-categories. The

[12] For simplicity, we are assuming that each buyer has the same preferences $\boldsymbol{\alpha}$. We could replace this assumption by assuming that the preferences $\boldsymbol{\alpha}$ of the buyers follow a normal distribution and accordingly estimate a mixed logit model. However, this more involved approach does not deliver significantly different results.

[13] We use a logit discrete choice model to elicit the preferences of the buyers. The scale of the logit discrete choice model is determined by the variance of the error terms ϵ_j. Thus, for our empirical analysis we can no longer use the convenient normalization of the price coefficient ρ to -1.

Covariates in buyer's utility fct.	**Job category** "Moving"	"Painting"	"Car"	"Plumbing"
Bid amount (€100)	-2.17***	-2.00***	-1.54***	-1.73***
Nbr. of positive ratings	.16***	.23***	.26***	.21***
Nbr. of negative ratings	-.19***	-.26***	-.20*	-.21**
Nbr. of employees	-.06	.02	-.28**	.06
Distance (km)	-.04**	-.07**	-.07	-.02
Trade License	.05	.17	-.07	-.11
Master craftsman company	.01	-.14	-.14	.10
Senior journeyman company	.11	-.08	-.17	-.23
Engineer	-.43	.28	-1.93*	.18
Technician	1.01	1.25*	-.56	-.10
Craftsman card	-.43	-.17	.08	.14
In craftsmen register	.07	-.05	.03	-.01
Certified registrations	.04	-.43	.53	.29
Other certifications	-.12	.03	-.04	-.17
Liability insurance	.43***	.06	-.02	.28
Certified membership	-.04	.10	.39**	.06
Constant	1.79***	1.65***	1.03***	1.01**
Nbr. of observations	12,161	6,119	2,614	2,435
Nbr. of auctions	2,599	1,140	702	606

Table II.2: **Preference estimates for startprice-category 1 and all job-categories.** The table gives the results of estimations of the logit discrete choice model given by (II.10) for startprice-category 1 and all job-categories. Displayed are the coefficients on the covariates in the utility functions of the buyers. Significance niveaus are reported by stars: ***: 1%, **: 5%, *: 10%.

results for all other job-startprice-categories are similar and due to reasons of brevity not displayed here.

The estimates for all job-startprice-categories exhibit the same general pattern: The coefficients on the price coefficient, the ratings coefficients and the constant are highly significant, while the coefficients on the other covariates are mostly insignificant. That does not come as a surprise, as the information about bidders most prominently displayed in the auction overview screen are bidders' prices and the number of their positive and negative ratings. Information on bidders' other non-price characteristics like the possession of certain degrees or the membership in certain institutions is only available after some additional clicks. The constant is highly significant because in about half of all auctions buyers choose to withdraw their job offers. It holds that the higher the value of the constant (which appears in the utility a buyer derives from a certain bidder), the lower is the value of the outside option.

Covariates in buyer's utility fct.	**Startprice category**			
	1	**2**	**3**	**5**
Bid amount (€ 100)	-2.17***	-1.42***	-1.32***	-.67***
Nbr. of positive ratings	.16***	.21***	.24***	.28***
Nbr. of negative ratings	-.19***	-.14***	-.22***	-.28***
Nbr. of employees	-.06	-.05	-.03	.02
Distance (km)	-.04**	-.04*	-.07***	-.12***
Trade License	.05	-.03	.12	.02
Master craftsman company	.01	.33	-.71*	-.86**
Senior journeyman company	.11	.25	-.25	-.43
Engineer	-.43	.24	.02	-.09
Technician	1.01	-.03	1.31	1.85***
Craftsman card	-.43	-.24	.12	.15
In craftsmen register	.07	-.09	-.07	-.26
Certified registrations	.04	-.37	-.08	.47
Other certifications	-.12	-.25***	.10	.18*
Liability insurance	.43***	.11	.70**	.76*
Certified membership	-.04	.26***	.03	.01
Constant	1.79***	2.02***	2.41***	1.16**
Nbr. of observations	12,161	9,037	8,092	8,573
Nbr. of auctions	2,599	1,665	1,358	1,209

Table II.3: **Preference estimates for job-category "moving" and all startprice-categories.** The table gives the results of estimations of the logit discrete choice model given by (II.10) for the job-category "moving" and all startprice-categories. Displayed are the coefficients on the covariates in the utility functions of the buyers. Significance niveaus are reported by stars: ***: 1%, **: 5%, *: 10%.

The numbers given in tables II.2 and II.3 are coefficient estimates and as such not directly interpretable. In order to get an impression of the effect of a decrease of a bidder's price by € 10 or an increase in his positive or negative ratings, we computed average marginal effects. For startprice-category 1 (table II.2) we find that a decrease of a bidder's price by € 10 increases his winning probability by around 2%. This holds for all job-categories. Over all job-categories, one additional positive rating increases a bidder's winning probability by around 1%, while an additional negative rating decreases a bidder's winning probability by around 2%. The influence of the number of ratings is most pronounced for the "plumbing" category, where one additional negative rating lowers a bidder's winning probability by even 3%.

For the job-category "moving" (table II.3), with respect to ratings we get the result that for all startprice-categories an additional positive rating increases a bidder's winning probability by around 1%, while an additional negative rating decreases a bidder's winning probability by around 2%. As might be expected, we find that the effect of a decrease in a bidder's price depends on the value of the auction (as is proxied for by the startprice) - the higher the value of the auction, the lower the effect of a certain price decrease. In particular, we find that while a price decrease of €10 increases a bidder's winning probability by 2% for startprice-category 1, it only increases a bidder's winning probability by 0.5% for startprice category 5.

We think it is reasonable to assume that on average jobs from the categories "moving" and "painting" require less skills than jobs from the categories "car" and "plumbing and heating". That is, for the latter categories we expect buyers to put more weight on the qualifications of bidders. This presumption is confirmed by our results - a look at table II.3 shows that the influence of a bidder's ratings relative to his price (as expressed by the relationship between the coefficient on a bidder's positive respectively negative ratings and the price coefficient) is indeed significantly higher for the categories "car" and "plumbing" than for the categories "moving" and "painting".

The results discussed above hinge on the assumption that the error terms ϵ_{nj} in (II.10) are neither correlated with the prices p_{nj} nor with the bidders' attributes $\mathbf{A}_{nj}$. In other words, for our estimation results to be consistent there must be no unobserved factors which influence buyers' utilities in a way systematically connected to our observables. However, as we analyze auctions conducted on an online marketplace, and as we were provided with very detailed recordings of these auctions, we are convinced that we are able to control for all factors which have a systematic influence on the buyers' utilities: Our data contains exactly the amount of information about bidders buyers have available when making their decisions. Thus, there should be no influences on buyers' utilities which are both unobserved and in some way systematically connected to bidders' attributes.

II.5 Analysis of Bidders' Information State

In section II.2 we proposed two models to describe bidders' behavior in open non-binding auctions. On the auction platform we have our data from bidders are informed about each other's non-price characteristics. Thus, their behavior should be in line with the predictions of our information case model. To verify this hypothesis, in this section we use a reduced form model to check whether the observed behavior of the bidders is indeed in line with the predictions of our information case model. In particular, we exploit contrasting testable predictions of both the information case and the no information case framework: If bidders behave according to our information case model, they should react to changes in the quality composition of an auction. In particular, in reaction to the appearance of a high quality opponent they should strongly decrease their prices. In contrast, if bidders behave according to our no information case model, they should show no reaction at all.

■ **Econometric model.** We test for these contrasting implications by using the following reduced form model of bidders' pricing behavior:

$$p_{nj} = \boldsymbol{\xi}\mathbf{K}_{nj} + \beta S_{nj} + a_j + \nu_{nj}. \tag{II.11}$$

This model describes bidders' pricing behavior along the lines of our theoretical frameworks from section II.2. Basically, we assume that the price bidder j puts forward in auction n depends on his costs c_{nj} and, in case of disclosed quality information, on his quality relative to that of his rivals. We assume the costs c_{nj} to depend both on the observable cost factors $\mathbf{K}_{nj}$ and on the unobserved opportunity costs of bidder j. How bidder j fares in terms of the buyer's valuation of his non-price characteristics (that is, in terms of quality) relative to his rivals is assumed to depend on bidder j's strength in terms of quality relative to the whole population of bidders and an unobserved auction-specific deviation. Bidder j's overall strength in terms of quality is captured in the bidder specific constant a_j. The error term ν_{nj} captures bidder j's opportunity cost for the job offered in auction n and the auction-specific deviation to this "overall strength".

The binary variable S_{nj} indicates whether bidder j has to face a rival bidder who is strong in terms of quality. We know from our theoretical considerations that if in case of disclosed quality information a rival of bidder j is replaced by one who is stronger in terms of quality, bidder j should react with a decrease in his price. In contrast, if quality information is concealed, bidder j should show no reaction. That means we expect $\beta < 0$ if bidders behave according to our information case model, and $\beta = 0$ otherwise.

■ **Identification strategy.** We restrict our analysis to bidders which are observed to participate in several auctions. In doing so, we are able to estimate equation (II.11) by mean-differencing (that is, employing a fixed effects estimator). By that we get rid of the individual specific and unobserved constants a_j. The assumption which has to hold for our estimates to be consistent is that the ϵ_{nj} are mean-independent from the observable cost elements $\mathbf{K}_{nj}$ and the strong rival indicator

S_{nj}. As we will discuss in more detail below, this assumption is likely to hold in our case.

■ **Estimation.** Our results from section II.4 show that throughout all job-startprice-categories the decisions of the buyers are strongly influenced by the number of positive and negative ratings bidders have. Thus, we define that a given bidder j encounters a strong rival in auction n if at least one of the other bidders in auction n has a difference of positive and negative ratings of at least 90:[14]

$$S_{nj} = \begin{cases} 1 & \text{if encounter with strong bidder (ratings difference} \geq 90\text{),} \\ 0 & \text{otherwise.} \end{cases}$$

In order to estimate equation (II.11) we need information about cost factors $\mathbf{K}_{nj}$. Thus, we have to restrict our estimation to auctions from the job category "painting". In section II.4 we saw that the number of positive and negative ratings a bidder has is of similarly strong influence on buyers' decisions over all startprice-categories. This allows us to use auctions from all startprice-categories for the estimations in this section. We want to estimate equation (II.11) by a fixed effects estimator, which means that we have to restrict our sample to bidders which are observed in at least two auctions. This leaves us with a sample of 941 bidders, taking part in 1,498 auctions from job category "painting" (the mean number of auction participations is 10, the median number is 6). In 22.2% of these auctions a bidder with a ratings difference of at least 90 takes part.

Table II.4 shows our estimation results. The first column displays our base specification. In column two we add dummies to control for auction size and for regional influences.[15] The coefficients on the cost factors do not vary much between the specifications, and they are of reasonable size: A professional tradesman in Germany charges on average €5-6 per painted square meter. This includes painting, paint, cleaning and travel. The average area to be painted in our data set is 138.3 m^2, the average travel distance 45.0 km (one-way). Together with our estimation results in table II.4, this implies that the average price per square meter painted, including paint and travel, is about €3-4 on the auction platform. Given that most of the bidders on the platform are non-professionals,[16] this number seems to be plausible. In both specifications the coefficient on the strong rival indicator S_{nj} is

[14]For comparison: The mean difference of positive and negative ratings in our sample is 5.8. 1% of the bidders in our sample have a ratings difference of at least 90.

[15]We define auctions to be from the same region when the first digit of their zip code is identical.

[16]78% of the bidders in our sample are neither master craftsmen nor senior journeymen.

Dependent variable: Bid amount of bidder j in auction n	(1)	(2)	(3)
Encounter with strong rival (dummy)	-82.85*** (14.89)	-91.57*** (17.73)	-93.79*** (17.77)
Controls:			
Area to paint (m^2)	1.72***	1.74***	1.61***
Area to paper (m^2)	1.41***	1.28***	1.29***
Paper removal (m^2)	2.72***	2.89***	2.54***
Cleaning (dummy)	77.63***	64.08**	54.01*
Reparation (dummy)	40.60***	56.39***	42.30***
Priming (dummy)	124.60***	125.44***	114.41***
No. of windows	10.41	11.00	13.39
No. of window frames	34.31	25.40	19.41
No. of doors	45.78***	46.23***	42.22***
No. of door frames	17.72***	18.56***	18.21***
Nbr. of radiators	85.33***	85.58***	78.91***
Paint by contractor (dummy)	25.99**	14.97	18.89*
Varnish by contractor (dummy)	125.58*	116.82	102.01
Distance (km)	1.15***	1.17***	.76***
Dummies for nbr. of bidders		X	X
Dummies for region		X	X
Controls for bidder composition			X
Bidder FE's	X	X	X
R^2	0.292	0.300	0.341
N	9,546	9,546	9,546

Table II.4: **Identification of the bidders' reaction to a strong rival; results of fixed effects estimation.** The table shows the results of a fixed effects estimation of the reduced-form model (II.11). The dependent variable is bid amount. Covariates are a dummy indicating the appearance of strong rival (a rival with a difference between positive and negative ratings of at least 90) and costs controls. The panel consists of 941 bidders who on average take part in 10 auctions each. Cluster-robust standard errors are reported in parentheses. For all results: both within- and between-R^2 are close to the overall R^2. Significance niveaus are reported by stars: ***: 1%, **: 5%, *: 10%.

highly significant and strongly negative, meaning that bidders bid more competitive if they encounter a strong rival: they lower their bids by around €90, which is a quite strong reduction if one considers that the average bid amount in our sample is about €550.

■ **Discussion of estimation results.** Our estimation results suggest that bidders react to the appearance of a strong rival by lowering their bids. This verifies our assumption that bidders behave according to our information case model. However, as mentioned during the derivation of equation (II.11) above, the coefficient at the strong rival indicator S_{nj}, β, can only be interpreted as the direct causal effect of the appearance of a strong rival on bidder j's bidding behavior if the unobserved part of equation (II.11), ν_{nj}, is mean independent from the observables $\mathbf{K}_{nj}$ and S_{nj}. ν_{nj} captures two unobserved influences on bidder j's bid: One stems from the composition of auction n in terms of the qualities of bidder j's rivals, the other stems from bidder j's cost components.

It might be that either strong bidders select themselves into certain auctions, or that certain types of bidders select themselves into auctions where a strong bidder is present. In effect, that would lead to a correlation between the appearance of a strong bidder and the auction-specific composition in terms of bidders' qualities. To be sure that we actually capture the bidders reaction to the appearance of a strong rival, in column 3 of table II.4 we control for the bidder composition of the different auctions. We do so by taking the averages over the attributes of all "weak" bidders (bidders with a difference of positive and negative ratings of less than 90), and using these averages as further controls in our fixed effects regression. As can be seen, controlling for the auction composition does not change our results. In addition, a large difference in positive and negative ratings is not correlated with any other of a strong bidder's attributes. Also, besides the prices put forward, the most prominent information auction participants are given is their rivals' ratings. Thus, we are pretty sure we are capturing the bidders' reaction to their rivals' differences in positive and negative ratings.

In contrast, possibly problematic for the identification of the bidders' information state is correlation between the covariates and the unobserved part of equation $(II.11)$ which stems from bidders' cost components. If the unobserved deviation in bidders' costs from their expected value is systematically connected to the appearance of a strong rival, significance of β would no longer indicate that bidders are informed about their qualities. However, there are two reasons why we do not think that the appearance of a strong rival is correlated with unobserved cost factors: First, we collected our data by extracting cost information from the job offers as they were available to the bidders. It is quite unlikely that we systematically missed a factor which is observable to the bidders and which indicates a deviation in costs. Second, even if we missed a factor of this kind, it should be known to the buyers. Before an auction starts, the buyers announce a startprice. This startprice is announced for informational purposes, and it should be reasonable to assume that, when setting the startprices, besides at strategic considerations buyers orientate themselves at the costs of their job. So, if there is a cost factor which is unobserved by us as researchers but known to the buyers and bidders, this cost factor should be reflected in the level of the startprice. Auctions in which a strong rival appears actually do systematically differ from auctions in which there is no strong rival in terms of the startprice. However, auctions in which a strong rival appears do not have a lower, but a higher startprice, indicating that strong rivals select themselves into auctions which seem to be quite valuable relative to the observable costs elements. This kind of selection should work against the hypothetical effect of the appearance of a strong rival in the case of informed bidders. As we are still able to observe more competitive bidding when a strong rival appears, we are quite certain that the coefficient on S_{nj} identifies strategic bidding behavior.

To summarize, our results strongly indicate that bidders have information about their qualities and that they behave as implied by our model for the information case: If a strong rival appears, bidding behavior becomes far more competitive. The competitive effect of the appearance of a strong rival is highly significant and robust against several controls. It manifests itself by price decreases of around 16%.

II.6 Counterfactual Analysis

In this section we determine the impact of availability of quality information on the aggregate welfare of the buyers. In our data, information about bidders' non-price characteristics is publicly available, and bidders can infer information about the preferences of the buyers regarding their non-price characteristics from observing buyers' former decisions. Thus, bidders' behavior should be in line with the information case model we developed in section II.2. In section II.5 we verified this assumption. We are interested in how buyers' welfare would change if non-price information was not available to the bidders. That is, we are interested in a counterfactual scenario where bidders are informed about each other's prices but not about each other's non-price characteristics. The buyers on the other hand shall always be informed about all bidders' prices and non-price characteristics.

In order to calculate the change in buyers' welfare if quality information was concealed, we need information about bidders' counterfactual prices. With information about bidders' costs c_{nj} we could calculate these counterfactual prices by employing our no information case model. We do not have explicit cost information, but as observed bidders' behavior seems to be in line with our model for the case of

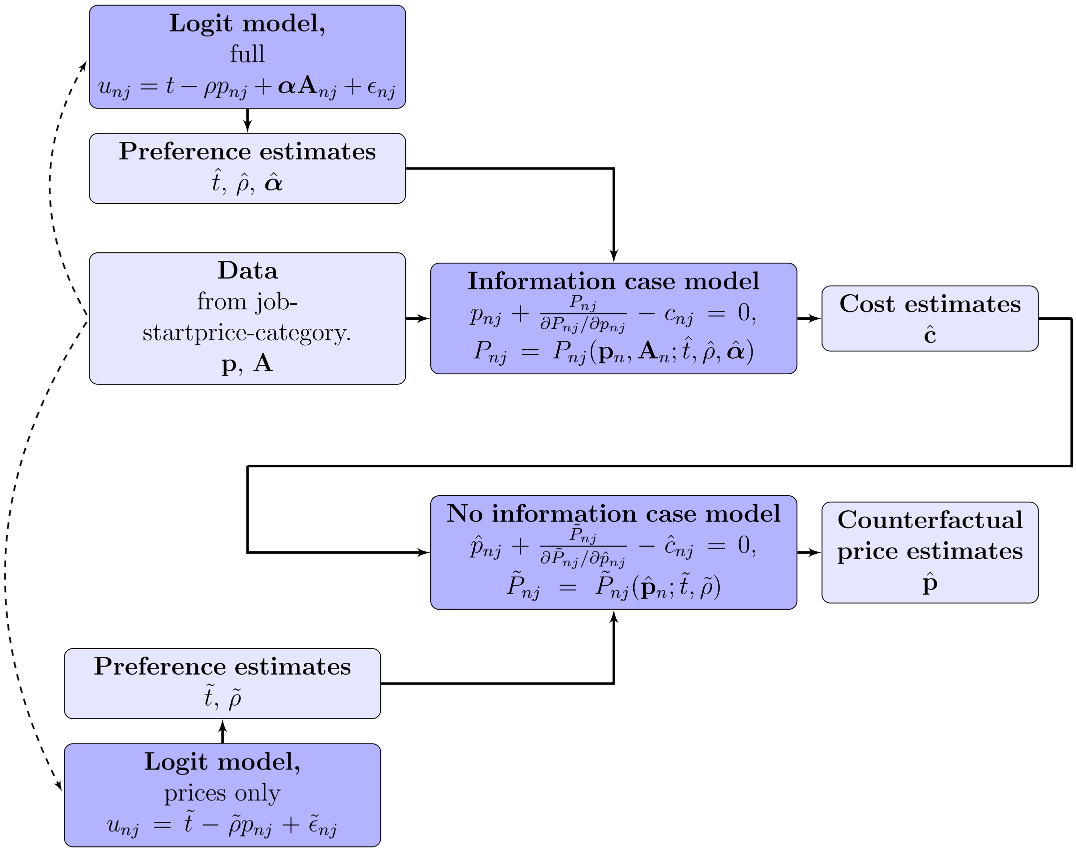

Figure II.4: **Sketch of the course of the counterfactual analysis.**

disclosed quality information, we can use this model to derive estimates of bidders' costs c_{nj} from the observed actual prices p_{nj}.

Our counterfactual analysis proceeds as follows: In our data we have information about bidders' prices and bidders' non-price characteristics. We use this information together with the information on buyers' preferences from section II.4 to solve our information case model (II.3) after estimates of bidders' costs $\hat{c}_{nj}$. We then use these cost estimates as input and solve our no information case model (II.6) after estimates of bidders' counterfactual prices $\hat{p}_{nj}$. Finally, we use our estimates of bidders' counterfactual prices $\hat{p}_{nj}$ to compute how buyers' welfare would change in case non-price information was concealed from the bidders. Figure II.4 depicts this course of our counterfactual analysis schematically.

■ **Estimation of bidders' costs.** Our assumption that bidders' behavior can be described by our model for the information case implies that the observed bids p_{nj} are equilibrium bids which for every auction n solve the bidders' first order

	Mean	*SD*	*Median*	*Mean*	*SD*	*Median*
	Moving			**Painting**		
Actual bidamounts (p_{nj})	€260.00	€84.36	€249	€256.80	€81.49	€248
Estimated costs ($\hat{c}_{nj}$)	€179.41 (€4.64)	€90.15	€167.25	€191.30 (€3.64)	€85.64	€179.65
Counterfactual bidamounts ($\hat{p}_{nj}$)	€263.80 (€2.02)	€83.67	€252.00	€266.83 (€1.97)	€81.04	€256.09
	Car			**Plumbing and Heating**		
Actual bidamounts (p_{nj})	€235.94	€74.73	€225	€233.38	€72.86	€224
Estimated costs ($\hat{c}_{nj}$)	€135.71 (€13.93)	€82.30	€123.64	€146.92 (€11.29)	€79.61	€136.93
Counterfactual bidamounts ($\hat{p}_{nj}$)	€235.58 (€4.52)	€74.48	€223.22	€242.74 (€6.50)	€71.97	€232.99

Table II.5: **Estimated costs and counterfactual bidamounts for startprice-category 2.** Displayed are summary statistics for the actual bidamounts, the estimated costs and the estimated counterfactual bidamounts for all four job categories and for startprice-category 2 (which includes startprices from €200-299). The results are based on 1,665 auctions for job category "moving", on 1,457 auctions for job category "painting", on 516 auctions for job category "car", and on 260 auctions for job category "plumbing and heating. Bootstrapped standard errors are given in parentheses.

conditions

$$p_{nj} + \frac{P_{nj}}{\partial P_{nj}/\partial p_{nj}} - c_{nj} = 0, \quad \forall j \in \{1, ..., J_n\}. \tag{II.12}$$

Besides on the bid amounts p_{nj} and the bidders non-price characteristics $\mathbf{A}_{nj}$, which we observe in our data, the winning probabilities P_{nj} depend on the preferences $\{\rho, \boldsymbol{\alpha}\}$ of the buyer. By inserting our estimates $\{\hat{\rho}, \hat{\boldsymbol{\alpha}}\}$ from section II.4, we directly arrive at estimates $\hat{P}_{nj}$ for the winning probabilities:

$$\hat{P}_{nj} = \begin{cases} \frac{1}{1+\sum_{k=1}^{J_n} e^{t+\hat{\rho}p_{nk}+\hat{\boldsymbol{\alpha}}\mathbf{A}_{nk}}} & \text{if } j = 0, \\ \frac{e^{t+\hat{\rho}p_{nj}+\hat{\boldsymbol{\alpha}}\mathbf{A}_{nj}}}{1+\sum_{k=1}^{J_n} e^{t+\hat{\rho}p_{nk}+\hat{\boldsymbol{\alpha}}\mathbf{A}_{nk}}} & \text{if } j \in \{1, ..., J_n\}. \end{cases} \tag{II.13}$$

With these, the first order conditions (II.12) can be solved for estimates $\widehat{c}_{nj}$ of the bidders' costs c_{nj}.

Table II.5 displays summary statistics of our cost estimates for startprice-category 2 and all four job-categories. To account for the fact that our cost estimates (as well as all other results of our counterfactual analysis) are based on estimates of

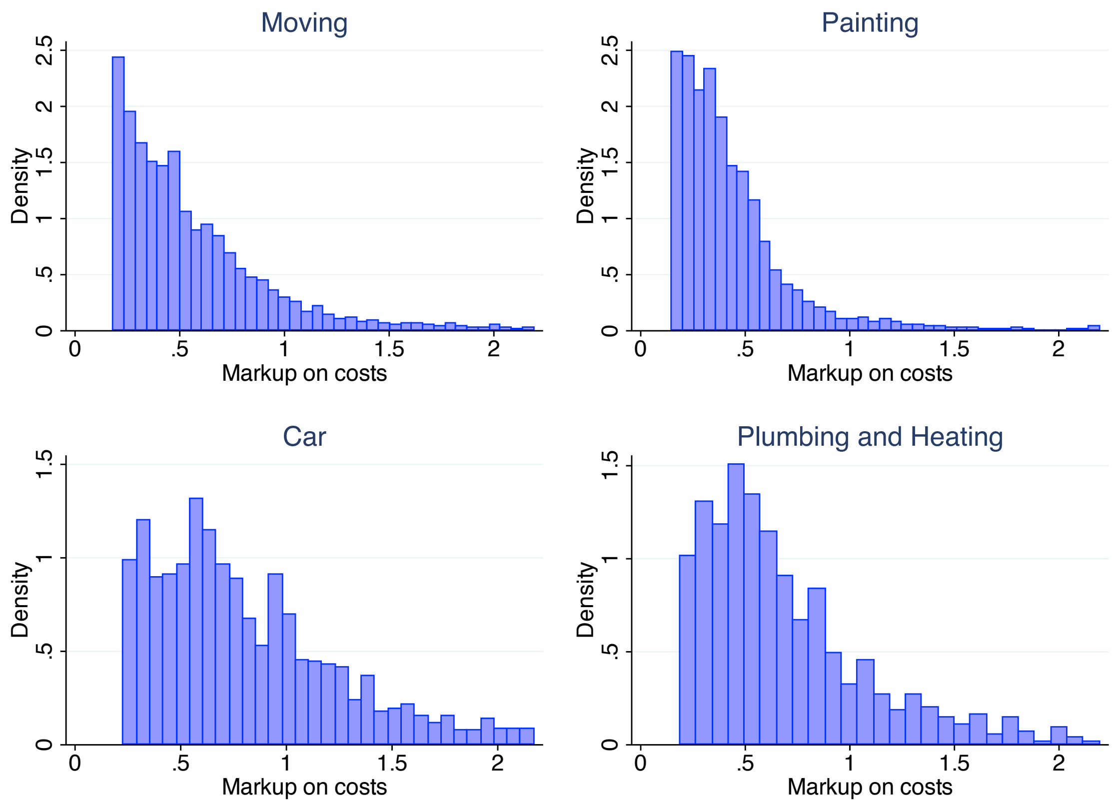

Figure II.5: **Distribution of bidders' markups.** Displayed is the density distribution of bidders' markups on their (estimated) costs for all four job categories and startprice-category 2 (which includes startprices from €200-299). Due to the sensitivity of our cost estimation to extreme bidamounts, for up to 5% of the bidders we get cost estimates close to zero and in thus in turn quite high markups. These are omitted here for the sake of illustration.

the buyers' preferences, we applied bootstrapping to get standard errors for our estimates. The standard error of the mean of our cost estimates ranges from €4-14. Thus, the estimates of bidders' costs are quite precise. The same holds true for the counterfactual results shown later on. The cost estimates become more meaningful if we look at the markup bidders demand on their costs. Figure II.5 displays the estimated distribution of bidders' markups on their costs for startprice-category 2 and all four job-categories. The median markup in the "moving" category is 47%, in the "painting" category it is 34%, in the "car" category it is 74%, and in the "plumbing" category it is 61%.

Now, are these markups of a sensible order of magnitude? From the cost information we manually collected for a part of the auctions from the "painting" category we know that for auctions from startprice-category 2 the average area to paint equals around 80 m^2. In more illustrative terms, that could mean, for example, to paint the walls and the ceilings of two small rooms of around 16 m^2 floor space each. We assume that, depending on the level of practice, a job like this could be done by one person in between four to eight hours. Startprice-category 2 includes auctions with startprices ranging from €200-299, and the level of bidders' prices is highly correlated with the level of the startprice. Given a markup of 30% to 40%, this would roughly amount to an hourly profit somewhere in the range of €10-20. Given that the median hourly wage before taxes in Germany is around €15 these seem to be plausible numbers.

As can be seen from figure II.5, for the job-categories "moving" and "painting" the majority of bidders demands markups of up to 50%. In comparison, in the job-categories "car" and "plumbing" the cost markups are significantly higher, with a major part of the bidders demanding markups between 50% and 100%. These results are in line with economic intuition: The qualifications required for jobs from the categories "car" and "plumbing" should on average be higher than that required for jobs from the categories "moving" and "painting". Thus, differences in qualifications among the bidders in the two former job-categories should be more pronounced than in the two latter categories, which in turn allows highly qualified bidders to demand larger markups in the categories "car" and "plumbing".

■ **Counterfactual Simulation.** Our counterfactual assumption is that non-price information is concealed from the bidders. In this case, the bidders' model of the buyers' decision process in a certain auction n is

$$\begin{aligned} \max_{j\in\{0,1,\dots,J_n\}} \quad & u_{nj}, \quad \text{where} \\ u_{n0} &= \tilde{\epsilon}_{n0}, && \text{(II.14)} \\ u_{nj} &= \tilde{t} - \tilde{\rho} p_{nj} + \tilde{\epsilon}_{nj} \qquad \text{for } j \in \{1,\dots,J_n\}. && \text{(II.15)} \end{aligned}$$

Like in the information case, also in the no information case we assume that bidders gather information about buyers' decision processes by observing past auctions. We can put ourselves in the bidders' position in the counterfactual no information case by ignoring the non-price information available to us as econometricians and estimating choice model (II.15) only using price information. With our estimates $\tilde{t}$ and $\tilde{\rho}$ we then can formulate the bidders' first order conditions in the no information case as

$$\hat{p}_{nj} + \frac{\tilde{P}_{nj}}{\partial \tilde{P}_{nj}/\partial \hat{p}_{nj}} - \hat{c}_{nj} = 0, \quad j \in \{1,\dots,J_n\}, \qquad \text{(II.16)}$$

where

$$\tilde{P}_{nj} = \frac{e^{\tilde{t}+\tilde{\rho}\hat{p}_{nj}}}{1+\sum_{k=1}^{J_n} e^{\tilde{t}+\tilde{\rho}\hat{p}_{nk}}}. \tag{II.17}$$

We solve conditions (II.16) numerically for estimates $\hat{p}_{nj}$ of bidders' equilibrium prices in the no information case.

With estimates $\hat{p}_{nj}$ of the counterfactual bids we can calculate the counterfactual aggregate utility of the buyers: Following Small and Rosen (1981), for type I extreme value distributed error terms ϵ_j the change in expected utility of the buyer in an auction n can be calculated as

$$\Delta\mathrm{EU}_n = \mathrm{EU}_n - \widetilde{\mathrm{EU}}_n = \ln\left(1+\sum_{j=1}^{J_n} e^{\hat{t}+\hat{\rho}p_{nj}+\hat{\boldsymbol{\alpha}}\mathbf{A}_{nj}}\right) - \ln\left(1+\sum_{j=1}^{J_n} e^{\hat{t}+\hat{\rho}\hat{p}_{nj}+\hat{\boldsymbol{\alpha}}\mathbf{A}_{nj}}\right).$$

The change in buyers' aggregate utility if quality information was concealed is then simply given as

$$\Delta\mathrm{EU}_{total} = \sum_{n=1}^{N} \Delta\mathrm{EU}_n \tag{II.18}$$

Division by $\hat{\rho}$ delivers the monetary equivalents of the changes in utility.

■ **Results.** Table II.6 displays the result of our counterfactual. For each job-startprice-category welfare changes are expressed in percentages of total revenues made (in monetary terms) in the respective category during the observation period. Total revenues range from around €180,000 in job-category "moving", startprice-category 1, to around €20,000 in job-category "plumbing", startprice-category 5. To account for uncertainty due to the fact that our results are based on estimates of the buyers' preferences, we computed bootstrapped standard errors. These are given in parentheses, together with the number of auctions on which the results for each job-startprice-category are based.

The changes in aggregate welfare of the buyers range from −8.6% in job-category "painting", startprice-category 5, to +8.7% in job-category "car", startprice-category 2. For each job-category there is a certain pattern of welfare changes: Roughly, for job-categories "moving" and "plumbing" the information structure does not considerably affect the aggregate welfare of the buyers. In contrast, for job-categories "painting" and "car" concealment of non-price information seems to have a clearly directed impact: While concealment of non-price information decreases buyers' welfare by up to around 9% for job-category "painting", it increases buyers' welfare

Startprice-category	**Job-category** Painting	Moving	Plumbing	Car
1 (€100-199)	-3.2% (1.2%, 1,140 auct.)	2.6% (0.8%, 2,599 auct.)	3.4% (3.4%, 606 auct.)	2.0% (3.4%, 702 auct.)
2 (€200-299)	-4.5% (0.9%, 1,457 auct.)	-0.1% (0.9%, 1,665 auct.)	-4.4% (4.1%, 260 auct.)	8.7% (6.5%, 516 auct.)
3 (€300-399)	-3.9% (0.9%, 1,302 auct.)	0.1% (0.9%, 1,358 auct.)	-1.0% (5.8%, 135 auct.)	4.6% (5.4%, 362 auct.)
5 (€500-599)	-8.6% (1.1%, 1,099 auct.)	-2.0% (1.1%, 1,209 auct.)	12.2% (17.4%, 119 auct.)	3.6% (23.6%, 179 auct.)

Table II.6: **Estimated changes in buyers' aggregate welfare in case non-price information gets concealed from the bidders.** For all job-startprice categories considered, the table displays the expected changes in buyers' welfare in case non-price information gets concealed. The percentage changes were derived by computing the monetary equivalent of the total change of buyers' welfare and then relating it to total auction turnover in the job-startprice-category considered. All auctions were conducted during the second half of 2008. The number of auctions and bootstrapped standard errors are given in parentheses.

by up to around 9% for job-category "car". With a look at the bootstrapped standard errors, for the "painting" category the results are significant throughout all startprice categories. For the "car" category the results are less pronounced, as the number of available observations is considerably lower there.

The pattern of welfare changes depicted in table II.6 can be explained along the lines of our considerations from section II.2. The level of skills required for jobs from categories "moving" and "painting" is lower than that required for jobs from categories "car" and "plumbing". Thus, for the two latter categories bidders should be more differentiated in terms of their qualities (that is, the buyers' valuations of their non-price characteristics) than for the two former categories. Hence for categories "moving" and "painting" we expect buyers' welfare to decrease when non-price information is concealed, while for categories "car" and "plumbing" we expect it to increase. The numbers in table II.6 show that for categories "painting" and "car" that is actually the case: When non-price information is concealed, for category "painting" we expect buyers' welfare to decrease by up to 9%, while for category "car" we expect it to increase by up to 9%. That we do not observe buyers' welfare to change in one clear direction for categories "moving" and "plumbing" can be explained by the fact that for these categories the relationship between bidders' costs and qualities is such that we are near the indifference line in figure II.1.

■ **Discussion.** The results of our counterfactual simulation are only meaningful if - although necessarily stylized - our theoretical framework captures the fundamental mechanics of the application at hand sufficiently well. We argue that this is the case: Our framework abstracts from inter-auction dynamics. That is, we assume that both buyers and bidders do not behave strategically across auctions. We think this assumption is reasonable for our application: First, as during the time period considered each buyer on average auctions off only one contract, we can exclude strategic inter-auction behavior of buyers. Second, the probability of repeated encounters between bidders is quite low: On average, a given bidder encounters only 12% of his rivals at least twice. Thus, it should be reasonable to assume that, if at all, phenomenons like tacit collusion play a negligible role. We also do not think that explicit collusion in a given auction plays a role: For once, bidders are not able to communicate with each other on the online platform. Then, as shown on the map in figure II.3, most auctions are procuring jobs in large cities respectively metropolitan areas. There, in contrast to rural areas, bidders should not know about the whole pool of potential rivals, what makes interactions between them apart from that on the platform unlikely.

A slightly different concern might be that some bidders behave strategically across auctions due to capacity constraints, like in for example Jofre-Bonet and Pesendorfer (2000). However, the auctions we consider are about smaller jobs which should take about one to at most three days to complete, and in the time span we consider (half a year) the average number of auction participations is around four. Thus, we do not think that capacity constraints do play a major role here. To summarize, we think that modeling each auction in an isolated manner is a reasonable approach for our application.

We further made the assumption that a bidding equilibrium emerges in each auction. In particular, this assumption implies that dynamic phenomenons like sniping do not occur in our application. Given the numbers in table II.1 this assumption seems to be justified: On average, the last bid is placed well before the end of an auction, meaning that sniping seems to play no role in our data. Thus, the assumption that in each auction in our application an equilibrium is achieved should be justified.

II.7 Conclusion

Non-binding reverse auctions are establishing as one of the most prominent tools for electronic procurement activities both for firms and government organizations. Whereas in non-binding auctions typically no structure is imposed on the buyer's decision process, important design questions arise, however, with respect to the information regime throughout the bidding process. We added to the understanding of this auction format by analyzing the effects of different designs of the information structure of an open non-binding auction. In particular, under the assumption that prices are always visible, we examined what effects disclosure respectively concealment of information about bidders' non-price characteristics has on the aggregate welfare of the buyers.

After establishing a formal framework, we first observed that buyers prefer that informational arrangement which creates higher competitive pressure among bidders. As we showed, which of the informational regimes indeed induces more competitive pressure crucially depends on the precise situation considered. Thus, from a theory point of view none of the regimes dominates.

To obtain further insights on the impact of the information regimes in non-binding auctions for real market situations, we then conducted an empirical analysis based on an extensive data set from a large European online procurement platform. The informational setup on this platform is such that bidders are informed about each other's non-price characteristics. Building on our formal framework, we performed a counterfactual welfare analysis to assess the consequences of concealing non-price information from the bidders. We find that our theoretical result - that the effect of concealment of non-price information depends on how strong buyers weigh bidders' non-price characteristics - also matters for applications in the field. In case buyers put a lot of weight on bidders' non-price characteristics, we find that if non-price information was concealed, buyers' welfare would increase by up to 9%. Contrary, if buyers put only low weight on bidders' non-price characteristics, we expect buyers' welfare to decrease by up to 9%. The latter is the case in the by far most popular job-categories on the platform.

The final policy recommendation implied by those results clearly depends very much on the final objectives of the online platform. Especially for business models in the very dynamic online markets, often rapid growth is much more important than instantaneous profits. In a recent interview for HBR IdeaCast from Harvard Business Review, Jeff Bezos, CEO of Amazon.com, for example states: “Percentage margins are not one of the things we are seeking to optimize. It's the absolute dollar-free cash flow per share that you want to maximize, [...]" And later on: "[W]e believe by keeping our prices very, very low, we earn trust with customers over time, and that that actually does maximize free cash flow over the long term." [17] A formal consideration of the dynamic aspects such as the long run profitability of firm growth in a specific sector by far exceeds the bounds of our structural analysis. Nevertheless, our analysis can contribute to questions arising in this broader context. If the most challenging task to achieve the long run growth objectives of the online platform indeed is to attract as many buyers as possible, then our results clearly show that the current information regime to reveal all non-price information is the one to best implement this objective, as it maximizes buyers' welfare in the most popular auction categories.

[17] Source: Interview with Jeff Bezos, HBR IdeaCast from Harvard Business Review, January 3, 2013.

Chapter III

Exploring the Opaqueness of the Patent System - Evidence from a Natural Experiment*

III.1 Introduction

The fundamental "deal" between the applicant of a patent and society is often stated as the granting of exclusion rights in exchange of the disclosure of a technical teaching that underpins the patented invention. In the assessment of the patent system's welfare balance, most work has focused on the incentive effect for inventors (respectively, applicants) - the "bait" of the promised exclusion rights is supposed to motivate inventors to spend more resources on research and development than they would in the absence of patent rights. The corresponding disclosure effects have received considerably less scrutiny, but many authors take as given that they exist and that they are sizeable.

Currently, the underlying assumption of policy makers and researchers alike is that the patent system is relatively transparent, that is, that searching for information on potentially conflicting prior art (which would limit the patentability of an invention) is rather costless. The patent system is apparently made to fulfill this ideal - disclosure by patent applicants is supposed to be complete, and insufficient disclosure can be held against the applicant by the examiner, leading to a refusal of the patent grant. At the same time, some users of the patent system have complained that the relevance of inventions is skillfully disguised by applicants who use arcane and

*This chapter is joint work with Dietmar Harhoff.

complex language in order to avoid in-depth scrutiny by rivals. Bessen and Meurer (2008), for example, explicitly recommend a reform which would require patent applicants to use "plain language" in order to avoid intransparent descriptions of the patented invention and excessive opaqueness of the overall patent system.

The evidence presented in this chapter supports the notion that the patent system is highly opaque. We use a quasi-experimental setting in which publicly available information - the request for acceleration - became private information at the EPO. Ex ante, the observability of the acceleration may have guided rivals of a patent applicant in detecting particularly valuable inventions. The information may have contributed to an above-average rate of oppositions against these patents (once granted).

We develop a theoretical model of the ex ante and ex post applicant and opponent behavior. Using aggregate data computed by the EPO from the non-public information on accelerations, we show that absent an observable signal of patent value, the likelihood of a patent being opposed drops sharply. The reduction in oppositions once the signal is no longer available suggests that potential opponents face problems in finding substitute signals or identifying the patent's contribution merely from the conventional data generated by the patent office. Given the quasi-experimental setting used here, we argue that our results provide fairly strong evidence in favor of the opaqueness presumption.

The chapter proceeds in four subsets. Section III.2 describes the institutional setting and thus the nature of the quasi-experiment. In section III.3, we specify a theoretical model in which requests for acceleration are related to patent value, and thus the likelihood of opposition. The model lends itself to developing a number of hypotheses regarding the identification and extent of opaqueness in the patent system. Section III.4 describes the empirical setup and provides a first set of descriptive results. Section III.5 concludes.

III.2 Institutional Background

The legal foundation for the activities of the European Patent Office (EPO) is given by the European Patent Convention (EPC) and a body of rules accompanying the Convention. The timing of patent filings and subsequent actions by the EPO is rather complex, but can be summarized as follows (see Harhoff and Wagner, 2009, for more details). Patent filings at the EPO are typically based on previous priority filings at national patent offices. These filings are then forwarded to the EPO and published there 18 months following the original priority date. With the publication of the patent document, the EPO also publishes a search report. The report is accessible to any third party and published prominently on the EPO's websites. Information contained in the search report may be crucial for the applicant in order to assess his or her chances of obtaining a strong patent grant. The information can also be crucial for rivals who wish to assess the legal strength of a particular patent.

Within six months of the publication, the patent applicant has to request the examination of the patent application, otherwise the patent filing will lapse. The examination process itself can be rather lengthy, and in many cases, applicants seek to delay the final decision by the EPO, since a patent grant with subsequent translation into national languages is rather expensive. However, some applicants are interested in fast resolution of patent examination. Reasons for being interested in a quick resolution may be that the patent holder wants to have the right to request an injunction against an infringer. Injunctions are only available after the patent application has been granted. Moreover, important investment decisions may have to be made by the applicant in order to enter product markets with patent protection. This may again explain why some applicants would like to see an acceleration of patent examination.

At the EPO, the applicant may request examination early (Article 96 (1) EPC). He may also unconditionally waive his right to receive an invitation from the EPO to confirm that he desires to proceed with the application. This waiver allows the application to reach the examining division more quickly. Typically, the request is made when filing the European patent application, but it can also be submitted later by separate communication to the EPO.

In Rule 93(d) - OJ2001, 458 - published on September 7, 2001 - the President of the European Patent Office announced that effective of December 3rd, 2001 (EPO 2001) requests for accelerated search and accelerated examination would no longer be made public (as they had been before).[1] This rule change meant that information that had been observable by any third party was now private information between the EPO and the applicant.

We exploit this setting by comparing statistics describing applicants' and their rivals' behaviors before and after the rule had been changed at the EPO. To have a

[1] Accelerated search is usually a precursor to accelerated examination. Therefore, we do not address any differences between the two institutions here. A differential treatment of the two proceedings is planned for an extended version of this chapter.

foundation for interpreting the statistics, we first develop a theoretical model in the subsequent section.

III.3 Model of Application and Opposition Process

We assume that there are two firms, firm A and firm B. Firm A shall have filed a patent application.[2] We further assume that firm A knows about the value of its patent. The value of firm A's patent is equivalent to the economic value of firm A's patent in case it gets granted. Besides by its value firm A's patent is characterized by its strength. The strength of the patent denotes the probability p that a patent is found valid in opposition cases. We assume that firm A uses the technology described in its patent to introduce an innovation into the market and that this innovation is rivalrous to firm B. That is, its introduction leads to a decrease in firm B's profits.

Firm A shall be able to choose between accelerated (a) and standard ($\neg a$) examination of its patent application. If firm A chooses accelerated patent examination and its patent gets granted, the profit firm A gains from its patented technology will be higher than in case of standard examination. Formally we will denote the profits firm A reaps from its patented technology by $\pi_{(\cdot)}^{(\cdot)}$. These profits shall depend on the value of the patent, which is either high (h) or low (l), and on whether the patent examination has been accelerated (a) or not ($\neg a$). Firm A's costs of accelerated patent examination shall be $c_a > 0$. We assume that firm A is only interested in accelerating a high-value patent: $\pi_h^a - \pi_h^{\neg a} > c_a$, and for simplicity $\pi_l^a = \pi_l^{\neg a} = \pi_l$. It shall hold that $\pi_h^a > \pi_h^{\neg a} > \pi_l$.

The present value of firm B's profits shall decrease if firm A's patent is granted. For simplicity we assume that firm A and firm B play a zero-sum game, which means that the gains of firm A in case of a patent grant equal firm B's losses.[3] That is, if firm A successfully patents its technology the profit $\pi(v, a)$ it makes equals the losses firm B incurs. To avoid reduction in its profits firm B has the possibility to oppose

[2] The assumption that the game starts after the patent application is filed simplifies the model, as the initial decision whether to file the patent application has not to be considered. That is, we consider the costs of filing the patent application to be sunk.

[3] For example, if firm A and firm B are the only players in a market of fixed size and firm A is able to increase its market share by a certain amount due to its patented technology, the market share of firm B decreases by the same amount.

the patent of firm A. If firm B decides to oppose firm A's patent both firms have to pay c_o for the then enfolding opposition process. At the end of the opposition process firm A's patent remains granted with probability p. This probability is the strength of firm A's patent.

The timing of the game firm A and firm B play is as follows: First, nature draws the value of firm A's patent, which can be high (h) or low (l). h shall be drawn with probability θ, l with probability $1 - \theta$. Then, firm A gets informed about the draw of the patent value, but firm B does not. After getting informed, firm A chooses whether to accelerate patent examination (a respectively $\neg a$). If firm A chooses accelerated patent examination it incurs costs c_a. Next, firm B has the possibility to oppose firm A's patent. In case firm B decides to oppose firm A's patent, both firms have to pay costs c_o. At the end of the then enfolding opposition process firm A's patent remains granted with probability p. Finally, payoffs are realized. These depend on the patent value (h respectively l) and whether patent examination has been accelerated (a respectively $\neg a$).

In the following we will differentiate between two information structures, which we call "public" and "private". In information structure "public", firm B is informed about whether firm A chose to accelerate patent examination, but firm B is not informed about the value of firm A's patent. In information structure "private", firm B is neither informed about firm A's acceleration decision nor about the value of firm A's patent. In other words, in information structure "public" there is a publicly visible signal related to patent value, whereas in information structure "private" this signal is concealed. The EPO's 2001 decision to make information about the applicant's acceleration decision no longer publicly available is equivalent to shutting down the value signal. Our main interest lies in a comparison of applicants' and rivals' behavior in information structures "public" and "private".

Before we sum up our game, we have to make some parametric assumptions. In general, if firm A's innovation is patented, firm A's profits increase and firm B's profits decrease. We already made the simplifying assumptions that we have a zero-sum game, which means that firm A's increase in profits equal firm B's decrease, and that acceleration is only worthwhile for high-value patents, which means that if a patent is of low value there is nothing to gain from acceleration. In order to establish a clear payoff-structure for our game, we make two small additional assumptions. First, we assume that the profits which can be gained from a low-value patent are larger than opposition costs and acceleration costs combined. Second, we assume that opposition costs are larger than acceleration costs.[4] Put together, all our parametrical assumptions are:

A1 We have a zero-sum game, that is firm A's gains from its patented innovation equal firm B' losses: $\pi_h^a(A) = \pi_h^a(B) = \pi_h^a$, $\pi_h^{\neg a}(A) = \pi_h^{\neg a}(B) = \pi_h^{\neg a}$, $\pi_l(A) = \pi_l(B) = \pi_l$.

A2 Firm A shall only be interested in accelerating a high value patent: $\pi_h^a - \pi_h^{\neg a} > c_a$, and for simplicity $\pi_l^{\neg a} = \pi_l^a = \pi_l$.

A3 The profit from a granted low-value patent is larger than opposition and acceleration costs combined: $\pi_l > c_o + c_a$.

A4 Costs of acceleration are smaller than costs of opposition: $c_a \leq c_o$.

Given these parametrical assumptions, the assumptions about the timing of the game, and the assumptions about the information structures, the extensive forms of our game are given by the game-trees in figures III.1 and III.2. Figure III.1 displays the extensive form of the game for information structure "public". First, nature draws the value of firm A's patent. With probability θ, the value is high (h), with probability $1-\theta$ it is low (l). Then, firm A gets informed about the value of its innovation in case it is patented and has to decide whether to accelerate its examination (a respectively $\neg a$). After firm A's acceleration decision firm B has to decide whether to oppose firm A's patent (o respectively $\neg o$). When making its decision, firm B is informed about whether firm A chose accelerated examination, but not about the value of firm A's innovation in case it is patented. The dashed lines in figure III.1 represent firm B's respective information sets: When firm B has to decide whether to oppose firm A's patent, it can base its decision only on information about whether firm A chose accelerated patent examination, but not on the actual value of firm A's patent.

Figure III.2 displays the extensive form of the game for information structure "private". Initially, the game proceeds as for information structure "public": Nature draws the value of firm A's patent, which is high (h) with probability θ and low (l) with probability $1-\theta$. Then firm A gets informed about nature's draw and, based

[4]These assumptions are made with a look at the field. Gambardella et al. (2008) estimate the median patent value to be €0.3 m. According to Levin and Levin (2002), opposition costs at the EPA amount to around €0.1 m. If a firm chooses accelerated patent examination it does not have to pay an extra fee but only to cope with increased administrative effort, which makes it sensible to assume that the costs of acceleration are smaller than those of opposition.

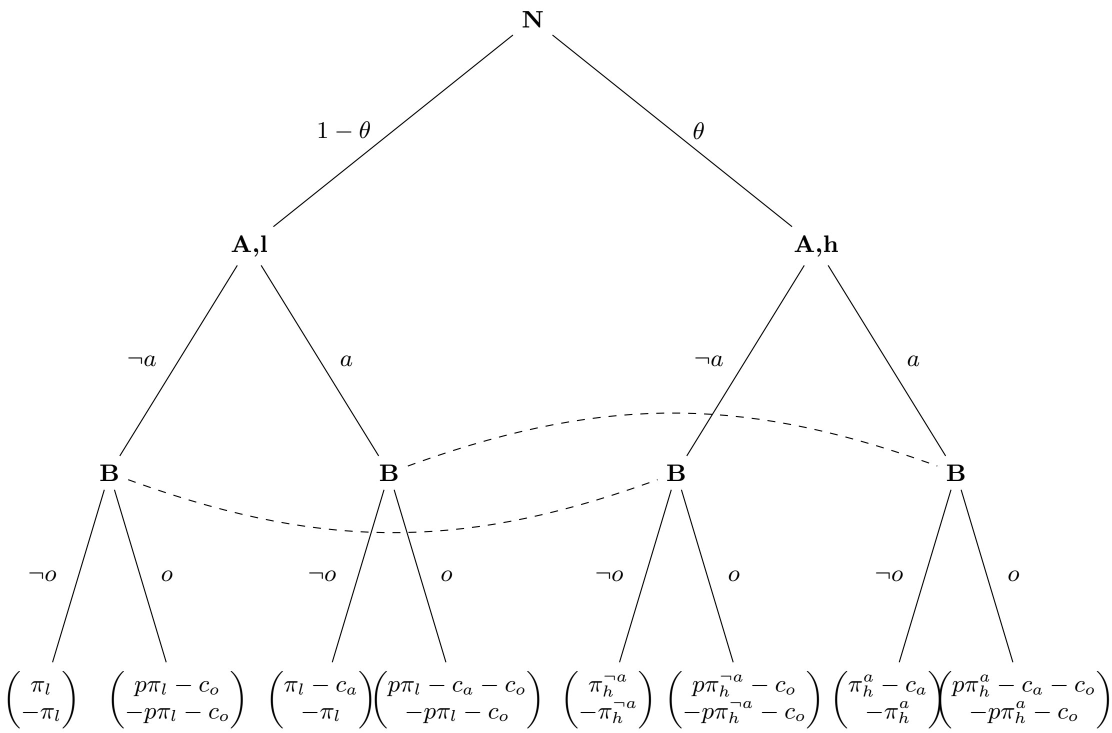

Figure III.1: **Extensive form of the game, information structure "public".** The graph shows the extensive form of the game firm A (the patent applicant) and firm B (its rival) play for the case that firm A's acceleration decision is disclosed to firm B. That is, whether firm A chooses accelerated patent examination (a respectively $\neg a$) is public information. The dashed lines mark firm B's information sets when it has to decide whether to oppose firm A's patent (o respectively $\neg o$).

on this information, decides whether to accelerate patent examination. In contrast to information structure "public", however, firm B is neither informed about the value of firm A's patent nor about firm A's acceleration decision. Accordingly, the dashed ellipse in figure III.2 represents the single information set of firm B: Firm B has to decide about opposing firm A's patent without information about both the value of firm A's patent and firm A's acceleration decision.

Solution

The game we set up above is a dynamic game of incomplete information.[5] We solve the game by applying the concept of the perfect Bayesian Nash equilibrium (PBNE). The PBNE is a modification of the more general concept of the Bayesian Nash equilibrium (BNE), which in turn is a modification of the most general Nash

[5] More specifically, for information structure "public" it is a signaling game in the tradition of Cho and Kreps (1987).

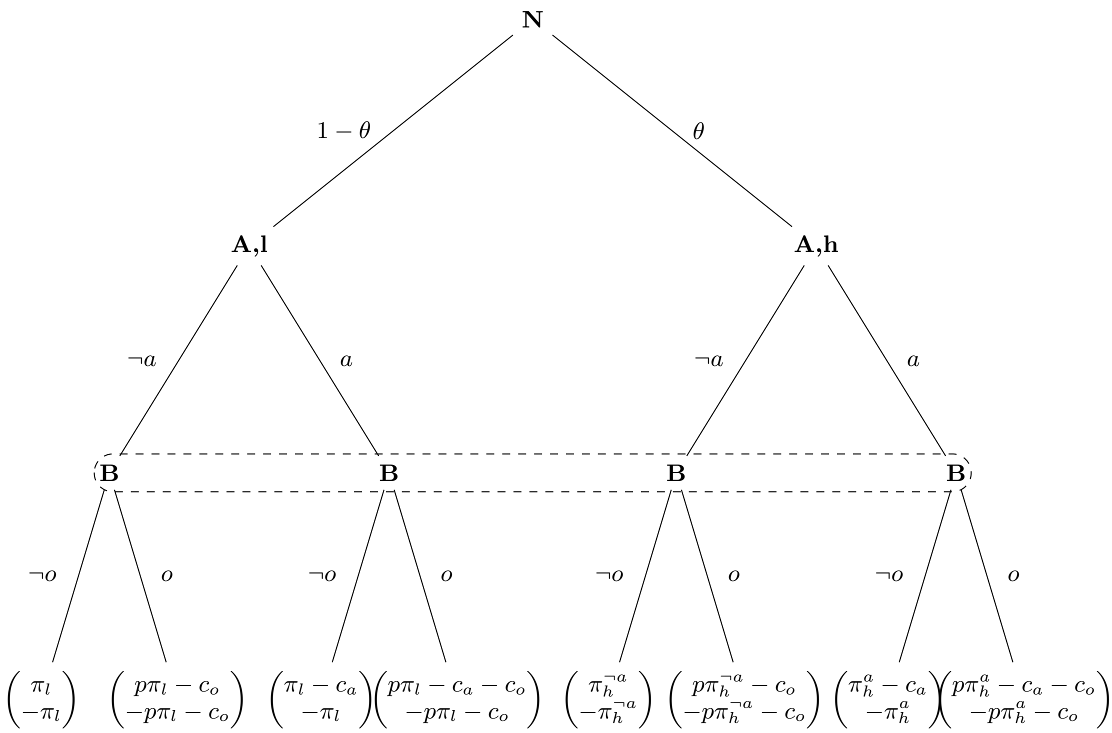

Figure III.2: **Extensive form of the game, information structure "private".** The graph shows the extensive form of the game firm A (the patent applicant) and firm B (its rival) play for the case that firm A's acceleration decision is concealed from firm B. That is, whether firm A chooses accelerated patent examination (a respectively $\neg a$) is private information to firm A. The dashed ellipse marks firm B's information set when it has to decide whether to oppose firm A's patent (o respectively $\neg o$).

equilibrium concept. The PBNE was introduced into game theory in order to rule out implausible equilibria in dynamic games of incomplete information, and exactly this is the purpose of its application here.

Figures III.1 and III.2 show that both for information structure "public" and "private" we have information sets with several nodes. If an information set contains several nodes, then the respective player has the same set of actions at every node of his information set. However, the player does not know at which node of the information set he actually is, but he has to form beliefs about his position inside the information set. A PBNE now demands two things: First, the strategies of each player have to be consistent with each player's beliefs. Second, each player obtains his beliefs from the equilibrium strategies and the observed actions by application of Bayes' rule. That is, strategies have to be consistent with beliefs, and beliefs have to be consistent with strategies.

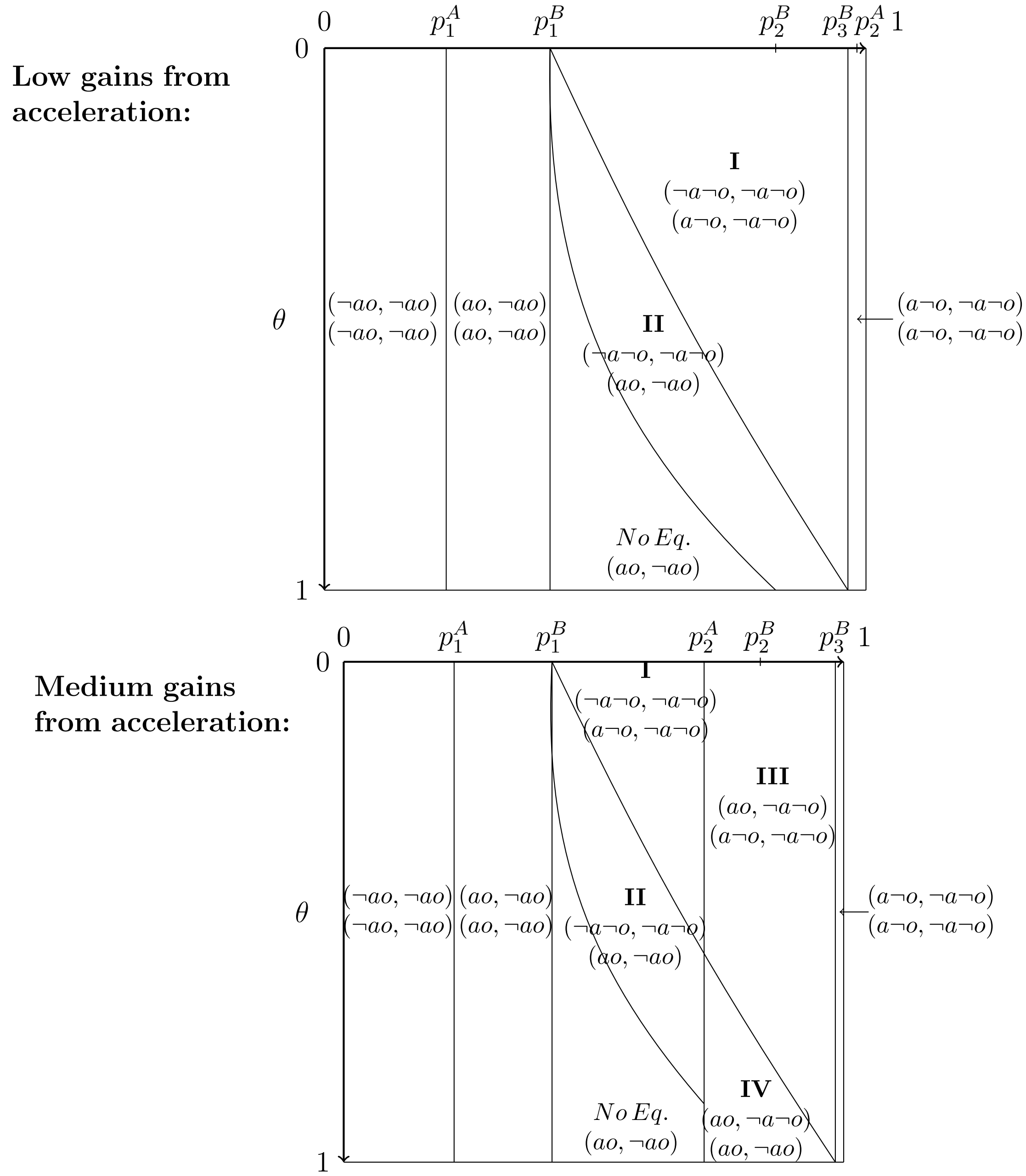

Figure III.3: **Outcomes for information structures "public" and "private" and low to medium gains from acceleration.** The upper graph displays the outcome structure in the p-θ-space for low gains from acceleration (that is, values of $\pi_h^a \in \Pi_3$), the lower graph for medium gains (that is, for values of $\pi_h^a \in \Pi_4 \cup \Pi_5$). For each subset of the p-θ-space, the groups of parentheses show the outcomes for information structures "public" (above) and "private" (below). In each parentheses, first the outcome in case firm A has a high-value patent and then the outcome in case firm A has a low-value patent is given. An outcome is described by firm A's action, that is acceleration (a) or no acceleration ($\neg a$), and firm B's subsequent action, that is opposition (o) or no opposition ($\neg o$).

High gains from acceleration:

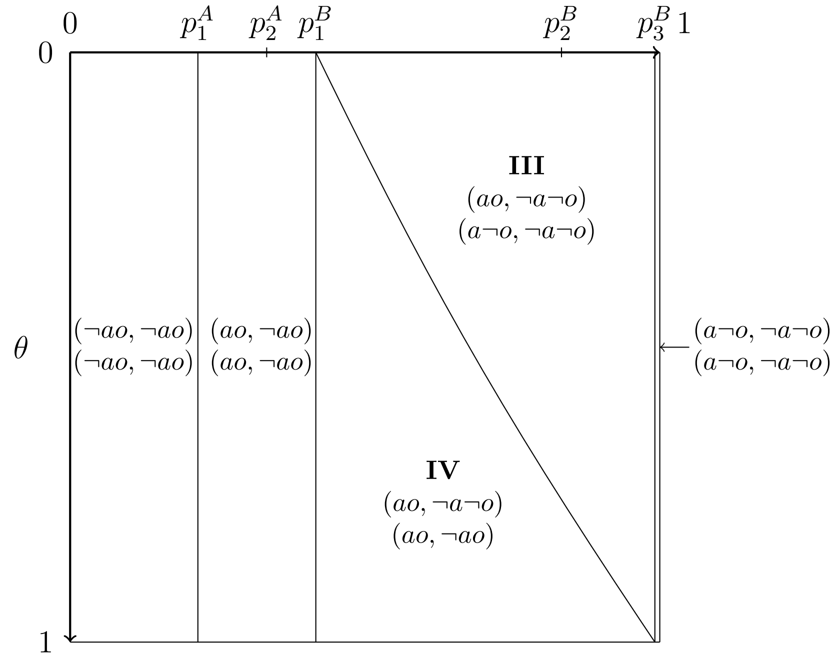

Figure III.4: **Outcomes for information structures "public" and "private" and high gains from acceleration.** The graph displays the outcome structure in the p-θ-space for high gains from acceleration (that is, for values of $\pi_h^a \in \Pi_6$). For each subset of the p-θ-space, the groups of parentheses show the outcomes for information structures "public" (above) and "private" (below). In each parentheses, first the outcome in case firm A has a high-value patent and then the outcome in case firm A has a low-value patent is given. An outcome is described by firm A's action, that is acceleration (a) or no acceleration ($\neg a$), and firm B's subsequent action, that is opposition (o) or no opposition ($\neg o$).

We apply the PBNE concept by first deriving BNE from the normal form of our game. For each BNE we then check whether it fulfills the criteria of a PBNE - that is, whether there is a belief structure which is consistent with this equilibrium. In doing so we will concentrate on equilibria in pure strategies. In addition, in order to rule out implausible equilibria we apply the "intuitive criterion", which was introduced in the context of signaling games by Cho and Kreps (1987). Roughly put, the intuitive criterion eliminates an equilibrium as "implausible" if a player using forward induction finds he would be better off if he deviated from that equilibrium.

In the end, the determination of the equilibria of our game hinges on payoff comparisons, and these payoff comparisons in turn depend on the specific relationships among our model parameters. In the following we structure the parameter space along three dimensions: The strength of the patent, that is the probability p with which the patent is found to be valid in case of opposition, the probability θ with which the patent is of high value, and the profit π_h^a firm A can gain from acceleration of a high-value patent. Given assumptions A1 to A4 above, the game can be solved for every subset of this parameter space. Appendix C.1 describes the solution of our signaling game in detail.

Results

We determined the equilibria of our signaling game for information structures "public" and "private" for all subsets of the parameter space. The parameter space is spanned along three dimensions: π_h^a, p and θ. Parameter π_h^a denotes the profit firm A gains from accelerating a high-value patent and spans from $\pi_h^a = \pi_h^{\neg a} + c_a$ (compare assumption A2) to infinity. Parameter p is the strength of firm A's patent, which is the probability that firm A's patent is found to be valid in case of opposition, and parameter θ is the probability that firm A's patent is of high value. Both parameters range from from zero to one.

Figures III.3 and III.4 display cross-sections through the π_h^a-p-θ space perpendicular to the π_h^a-axis. That is, each of the three depicted planes corresponds to one particular value of π_h^a. As figures III.3 and III.4 show, these cross-sections are further divided into subsets by cut-off values $p_{(\cdot)}^{(\cdot)}$. The ordering of these cut-off values depends on the value of π_h^a at which a cross-section was produced. With respect to the position of the cut-off values relative to each other we can divide the π_h^a space into several subsets. These subsets are described in detail in appendix C.1. In figures III.3 and III.4 only representative cross-sections for subsets of the π_h^a space which are associated with substantial gains from acceleration are displayed.

■ **Outcome patterns.** We find that only for patents of intermediate strength our results do depend on the information structure. For the discussion of our results we structure the subset containing patents of intermediate strength ($p_1^B < p < p_3^B$) further into subsets I to IV.[6] These subsets are marked in figures III.3 and III.4. Our results with respect to firms' behavior are summarized in the following proposition:

Proposition 1 (Outcome patterns.). *The outcome patterns for weak ($0 < p < p_1^B$) and strong ($p_3^B < p < 1$) patents do not depend on the information structure. For patents of intermediate strength ($p_1^B < p < p_3^B$), the behavior of the firms depends on the information structure:*

[6] There is no Nash equilibrium and thus also no PBNE for one of the subsets which contain patents of intermediate strength. In the following discussion we therefore will ignore this subset. As this subset describes the quite unrealistic situation that high-value patents occur with very high probability, we do not lose information which is of economic significance.

*i) Outcome patterns in case information about firm A's acceleration decision is **concealed**:*

Firm A does not accelerate low-value patents but accelerates high-value patents. If high-value patents are rare (subsets I and III), firm B does not oppose. If high-value patents are more frequent (subsets II and IV), firm B opposes.

*ii) Outcome patterns in case information about firm A's acceleration decision is **disclosed**:*

If the patent is rather weak and gains from acceleration are rather low (subsets I and II), firm A never accelerates and firm B never opposes.

In contrast, if the patent is rather strong and the gains from acceleration are high (subsets III and IV), then firm A accelerates high-value patents only and firm B opposes if it observes acceleration.

The derivation of the results in proposition 1 can be found in appendix C.1. Below we discuss the outcome patterns in firms' behavior.

The outcome patterns for weak and strong patents are the same for information structures "public" and "private". The reason is that for both weak and strong patents firm B's decision whether to oppose is not influenced by information possibly emitted by firm A's acceleration decision: If firm A's patent is weak ($0 < p < p_1^B$), then it is worthwhile for firm B to oppose firm A's patent regardless of its value. Accordingly, when deciding whether to accelerate the examination of its patent, firm A does not have to take into account the signaling effect of its decision - firm A simply chooses to accelerate patent examination when this decision makes it better off in expectation conditional on opposition by firm B. The same logic applies if a patent is very strong ($p_3^B < p < 1$): In this case firm B never gains from opposition, and again in its decision whether to accelerate patent examination firm A has only to take into account its payoffs (conditional on no opposition by firm B). In effect, for strong patents firm B never opposes and firm A accelerates only if it has a high-value patent. In contrast, outcome patterns for patents of intermediate strength ($p_1^B < p < p_3^B$) differ between the cases of concealed and disclosed acceleration information.

□ **Concealed acceleration information.** In case information about firm A's acceleration decision is *concealed* from firm B, for patents of intermediate strength (that is, for subsets I to IV) firm A always plays a separating strategy. Firm A's separating strategy in subsets I to IV is: Acceleration in case the patent is of high value, no acceleration in case it is of low value. In contrast to firm A, which plays the same strategy in all subsets, firm B plays different strategies in subsets I and III and subsets II and IV: In subsets I and III, which are related to low probabilities that a patent is of high value, firm B chooses not to oppose, whereas in subsets II and IV, which are related to high probabilities that a patent is of high value, it chooses to oppose.

We first explain why firm A plays a separating strategy in all subsets: As firm A's acceleration decision is concealed, by choosing to accelerate its patent it does not transmit any information to firm B which could possibly trigger opposition. Also, regardless of firm B's action, in the end firm A is always better off if it chose to accelerate a high-value patent. As firm A does not profit from accelerating low-value patents, it is thus best for firm A to play a separating strategy: Accelerate high-value patents, but do not accelerate low-value patents.

Let us now turn to firm B's reaction: In the upper subsets I and III the probability that a patent is of high value is low. That means it is far more likely that firm A has a low-value patent, which firm B would not want to oppose, than a high-value patent, which firm B would want to oppose. Thus, in the upper subsets I and III firm B is better off in expectation if it does not oppose firm A's patent. In contrast, in subsets II and IV the probability that a patent is of high-value is relatively high. That is, in subsets II and IV it is far more likely that a given patent is of high-value (and accelerated) than that it is of low-value, and thus now firm B is better off in expectation if it opposes firm A's patent.

☐ **Disclosed acceleration information.** Let us first take a look at subsets I and II, for which the gains from acceleration of a high-value patent are rather small. For subset I and subset II the outcomes in case acceleration information is *public* are the same: Firm A never accelerates its patent, and firm B never opposes. The reasoning is as follows: Firm B knows that acceleration is only worthwhile for a high-value patent. Thus, if it observed firm A to accelerate patent examination, firm B would conclude that firm A has a high-value patent. As firm B only gains from opposing a high-value patent, it would oppose if firm A accelerated its patent, which in turn would mean that firm A's expected profits from acceleration (which are small here) are offset by the costs of the then enfolding opposition process. Thus, firm A plays a pooling strategy - it chooses not to accelerate its patent regardless of its value. Given that firm A plays a pooling strategy, firm B cannot infer information about the value of the patent from firm A's action, and therefore has to base its decision whether to oppose solely on the prior probability of a high-value patent (which equals θ). For subsets I and II the probability that firm A's patent is of high value is rather small, and therefore firm B, which in subsets I and II can only profit from opposition of a high-value patent, is in expectation better off if it does not oppose firm A's patent.

Both in subset III and subset IV firm A accelerates high-value patents only and firm B only chooses to oppose when it observes acceleration. The reason is that in subsets III and IV additional profits of firm A from acceleration are so high that the expected gains from acceleration of a high-value patent outweigh the costs of an opposition process. Thus, firm A accepts that by accelerating it sends a signal which induces firm B to oppose, because in expectation firm A is better off if an accelerated high-value patent is opposed than if a not-accelerated high-value patent is not opposed. Firm B, which still only can profit from opposition of a high-value patent, reacts accordingly by opposing only when it observes acceleration.

■ **Partial welfare analysis.** We conduct a partial welfare analysis with respect to the applicant, firm A, and its rival, firm B. For all subsets of the parameter space we make welfare comparisons between information states "public" and "private". The following proposition summarizes our results:

Proposition 2 (Partial welfare analysis.)**.** *For weak ($0 < p < p_1^B$) and strong ($p_3^B < p < 1$) patents there is no change in firms' welfare between information structures "public" and "private". For patents of intermediate strength ($p_1^B < p < p_3^B$), the welfare of the firms depends on whether firm A's decision to accelerate its patent is concealed from firm B:*

*i) For **low probabilities that a patent is of high-value** (subsets I and III), firm A is better off in case acceleration information is concealed, whereas firm B is better off in case acceleration information is disclosed. If patents are weak (subset I), the aggregate welfare of firm A and firm B is highest in information structure "public". If patents are strong (subset III), it is highest in information structure "private".*

*ii) For **high probabilities that a patent is of high value** (subsets II and IV) firm A and firm B are better off in case acceleration information is concealed, both if considered individually and if considered together.*

Analytical welfare comparisons can be found in appendix C.2. Table III.1 summarizes the results. We find that for weak ($0 < p < p_1^B$) and strong ($p_3^B < p < 1$) patents the welfare of firm A and firm B is the same for both information structures. The reason is simply that there is no difference in outcomes between information structures (compare figures III.3 and III.4 and proposition III.3). Below we discuss our welfare results for patents of intermediate strength.

□ **Welfare - Low probabilities of high-value patents.** For patents of intermediate strength and for *low probabilities that a patent is of high value* (subsets I and III) the welfare results follow clear patterns: Firm A is the better off the less information is available to firm B about the value of firm A's patent. The intuition is pretty simple: The less information firm B receives about the value of firm A's patent, the more information rent firm A can extract - in a way, the less information about the value of a patent is available, the better firm A can "hide" its few high-value patents among the bulk of low-value patents. Conversely, with more information available about firm A's patent firm B becomes better off. The reason is that firm B can only profit from opposing high-value patents. Thus, the more information firm B has about the value of firm A's patent, the more targeted it can be in its opposition activities, and the less resources are wasted on low-value patents.

So, for low probabilities that a patent is of high value the interests of firm A and firm B diverge: While firm B prefers information structures which reveal information about the value of firm A's patent, firm A prefers information structures which conceal this information. Thus, which information structure is best with a look at the combined welfare of firm A and firm B depends on how large the gains of firm A from concealed value information are relative to the benefits of firm B from disclosed value information:

In subset I opposition by firm B is quite likely to be successful. As the probability of successful opposition is rather high, firm A only accelerates its patent in case information about its acceleration decision is concealed. Without an observable signal transporting value information firm B refrains from opposition, as firm B can only profit from opposition of a high-value patent, and as it is quite unlikely that a given patent is of high-value. From an aggregate welfare perspective the additional gains of firm A from acceleration of its high-value patent do not matter, as we made the assumption that the gains of firm A from the introduction of its patent equal the losses of firm B. Thus, in case the acceleration signal is concealed aggregate welfare

	Subset I	Subset III
Firm A	Private	Private
Firm B	Public	Public
A + B	Public	Private
	Subset II	**Subset IV**
Firm A	Public	Public
Firm B	Private/Public	Public
A + B	Public	Public

Table III.1: **Results of the partial welfare analysis.** For subsets I to IV the table displays what information structure results in the highest welfare for firms A and B if considered individually and together. The result for firm B in subset II depends on p: Firm B is better off with information structure "public" in case $p > p_W$. Otherwise, firm B is better off with information structure "private". The cut-off value p_W is given in appendix C.2.

is determined by the costs firm A directly incurs from accelerating its patent. In case the acceleration decision is public information the threat of being opposed by firm B lets firm A refrain from acceleration. Also, due to the low a-priori probability of a high-value patent firm B still does not oppose. Thus, in subset I combined welfare of firm A and firm B is highest for information structure "public", as there firm A does not incur acceleration costs.

In contrast to subset I, in subset III chances that opposition of firm B is successful are small. Thus, firm A now accelerates its high-value patents both in case its acceleration decision is disclosed and in case it is concealed. In case firm A's acceleration decision is concealed firm B refrains from opposition due to the low a-priori probability of a high-value patent. In contrast, in case firm A's acceleration decision is disclosed, firm B opposes if it observes acceleration. The probability that firm B succeeds with opposition is rather low, but both firm A and firm B incur opposition costs. That is, firm B's opposition effort in case the acceleration signal is disclosed is more or less "wasted", and therefore combined welfare of firm A and firm B is highest if information about firm A's acceleration decision is hidden from firm B ("private").

□ **Welfare - High probabilities of high-value patents.** For patents of intermediate strength ($p_1^B < p < p_3^B$) and for high probabilities that a patent is of high value (subsets II and IV) firm B becomes better off the more information it receives about the value of firm A's patent. The reason is that firm B only profits from opposition of a high-value patent, and thus more information about the value of the submitted patent allows firm B to use its resources more efficiently. However, firm A no longer becomes better off the less information about the value of its patent is available to firm B. To the contrary, it prefers information structures in which information about the value of its patent is disclosed to firm B. The reason is that without any discriminating information about patent value firm B opposes all patents, as because of the high a-priori probability of high-value patents the number of expected "hits" (opposition of a high-value patent) makes up for the "misses" (opposition of a low-value patent). Given a signal about patent value, firm B can use its resources more efficiently by restricting its opposition efforts to high-value patents, and firm A gets at least its low-value patents granted for sure. Thus, in subsets II and IV both firm A and firm B benefit when information about the value of firm A's patent is disclosed (information structure "public"). Of course that means that also the combined welfare of firm A and firm B is highest if information about the patent value is transmitted by the acceleration signal.

■ **Opaqueness - welfare implications.** Above, we maintained the assumption that the patent system is opaque with respect to the value of patents. Given that the patent system is opaque, we analyzed how firms' behavior and firms' welfare change in case information about the applicant's acceleration decision becomes concealed. Here, we compare an opaque patent system (with and without a publicly observable acceleration signal) to a transparent patent system where both the value of firm A's patent *and* its acceleration decision. That is, we turn our attention to the implications of opaqueness of the patent system. The following proposition summarizes our results:

Proposition 3 (Welfare implications of opaqueness of the patent system.)**.** *For weak ($0 < p < p_1^B$) and strong ($p_3^B < p < 1$) patents outcomes in case the patent system is opaque with regard to patent value are equal to outcomes in case the patent system is transparent. For patents of intermediate strength ($p_1^B < p < p_3^B$) the effect of opaqueness on the aggregate welfare of firm A and firm B depends on whether the probability of a high-value patent is low or high:*

*i) For **low probabilities that a patent is of high-value** (subsets I and III), aggregate welfare of firm A and firm B is (weakly) higher in case the patent system is opaque than in case it is transparent.*

*ii) For **high probabilities that a patent is of high value** (subsets II and IV), the effect of opaqueness on the aggregate welfare of firm A and firm B depends on whether in case the system is opaque firm A's acceleration decision is disclosed or concealed:*

In case firm A's acceleration decision is disclosed, aggregate welfare of firm A and firm B is (weakly) higher than in case the patent system is transparent.

In case firm A's acceleration decision is concealed, aggregate welfare of firm A and firm B is lower than in case the patent system is transparent.

The computation of the outcomes in case the patent system is transparent can be found in appendix C.1.

It is easy to see why for weak and strong patents the outcomes in case the patent system is transparent are equal to those in case the patent system is opaque: If patents are weak ($0 < p < p_1^B$) or strong ($p_3^B < p < 1$), firm B's opposition decision is independent from the value of firm A's patent. In case patents are weak, it is always worthwhile for firm B to oppose. On the other hand, in case patents are strong, it is never worthwhile for firm B to oppose. That is, information about patent value simply does not play a role and therefore outcomes in case the patent system is transparent are equal to those in case the system is opaque. However, for patents of intermediate strength ($p_1^B < p < p_3^B$) aggregate welfare of firm A and firm B depends on whether the probability of a high-value patent is low or high:

□ **Low probabilities of high-value patents.** For low probabilities that a patent is of high value (subsets I and III) aggregate welfare of firm A and firm B is (weakly) higher in case the patent system is opaque than in case it is transparent. In case of a transparent patent system the individual incentives are such that firm A accelerates high-value patents and firm B opposes high-value patents. In subset I patents are weak, whereas in subset III they are strong. If patents are weak, there is a high probability that in case of opposition the patent is revoked. Thus, in case of a transparent patent system acceleration efforts of firm A would be wasted on average. In case of an opaque patent system firm A either refrains from acceleration (information structure "public") or it accelerates its high-value patents "in disguise" (information structure "private"). In reaction, as high-value patents are hidden among the bulk of low-value patents, firm B refrains from opposition. In both cases there is no waste of acceleration costs, and therefore from an aggregate perspective firm A and firm B are better off with an opaque patent system.

In subset III patents are strong. That is, the probability that a patent remains granted in case of opposition is high. Thus, in case of a transparent patent system opposition efforts of firm B are wasted on average. If the patent system is opaque and acceleration information is disclosed, the outcome is the same as in case of a transparent patent system. If the patent system is opaque and acceleration information is concealed, firm B refrains from opposition as it cannot target high-value patents and the probability of a "hit" (that is, opposition of a high-value patent) is low. In this case there is no waste of opposition efforts (relative to the situation of a transparent patent system), and thus aggregate welfare of firms A and B is (weakly) higher with an opaque than with a transparent patent system.

□ **High probabilities of high-value patents.** For high probabilities that a patent is of high value (subsets II and IV), the effect of opaqueness on the aggregate welfare of firm A and firm B depends on whether in case the system is opaque firm A's acceleration decision is disclosed or concealed. In case the patent system is opaque and firm A's acceleration decision is disclosed, the outcome for strong patents (subset IV) is the same as for a transparent patent system. The outcome for weak patents (subset II) differs from that for a transparent patent system in that high-value patents are neither accelerated by firm A nor opposed by firm B. That is, acceleration costs are not wasted and thus aggregate welfare of firms A and B is higher in case the patent system is opaque and information about firm A's acceleration decision is disclosed than in case the patent system is transparent.

In case the patent system is opaque and firm A's acceleration decision is concealed, firm B opposes all patents. The reason is that the probability of a "hit" (that is, opposition of a high-value patent) is high in subsets II and IV. In contrast, in case of a transparent patent system firm B focuses its opposition efforts on high-value patents. As firm B does not profit from opposition of a low-value patent, in case the patent system is opaque and there is no acceleration signal opposition costs are inefficiently high. Thus, from an aggregate welfare perspective in subsets II and IV firms A and B are better off with a transparent patent system.

III.4 Empirical Evidence for Opaqueness of the Patent System

In December 2001, the EPO changed its information policy regarding acceleration requests of patent applicants: While before December 2001 information about acceleration requests of applicants was publicly available, after December 2001 this information was concealed from the public. If the European patent system is indeed opaque with respect to patent value, we expect this change in the EPO's information policy to impact the behavior of both patent applicants and their rivals. In this section, we use data provided to us by the EPO to look into empirical evidence on opaqueness of the European patent system with respect to patent value. We first use the theoretical framework developed above to structure our predictions about the way the EPO's 2001 policy change affected the behavior of the parties involved in the patent application process. We then take a look at the data to see how the behavior of applicants and rivals changed in reaction to the EPO's 2001 policy change. From the way behavior changes we can draw conclusions about whether the European patent system is opaque with respect to patent value. Also, we can give a first assessment of the welfare implications of the EPO's 2001 decision to conceal acceleration information.

Predictions from our model

In section III.3 we developed a model which captures the essential mechanics of the patent application and opposition process: We assumed that there are two possible types of patents (high-value and low-value), that each type occurs with a certain probability, that there are certain gains and costs from patent acceleration, and that opposition is costly and successful with a certain probability. The relationship between these parameters determines the predictions of our model regarding changes in the behavior of firm A (the applicant) and firm B (its rival) in case firm A's acceleration request gets concealed from the public. Instead of data on the behavior of single applicants and rivals we have available aggregate data on the fractions of yearly filings which were accelerated respectively opposed. Accordingly, when deriving predictions about the effects of the EPO's 2001 change in its information policy, in the following we will interpret changes in the outcomes of our model as changes in the respective frequencies which we observe in our data.

■ **Statements about opaqueness of the patent system.** Intuitively, one might think that if the patent system was fully transparent with regard to patent value we would observe no behavioral changes at all in reaction to concealment of information about the applicants' acceleration requests, and that we thus in turn could infer from the observation of changes in either the acceleration or the opposition frequency that the patent system is opaque. However, a closer look at our model reveals that this intuition is not exactly right: Acceleration of patent examination is assumed to increase the expected profits the submitting party can reap from its patent. If we make the assumption that what the submitting party gains when its patent is granted is at the expense of its rivals, then the incentive to oppose the patent application might increase if accelerated examination (which makes the patent more worthwhile for the submitting party and thus more hurtful for its rivals) is requested. Thus, even if the patent system was perfectly transparent with respect to the "base value" of a patent, we still might observe changes in the behavior of the involved parties in case the EPO changes its information policy and conceals the acceleration signal from rivals, simply because rivals are no longer informed about the "net value" of the patent.

We can only be sure that behavioral changes as a reaction to the 2001 change in the EPO's information policy are solely caused by the loss of a signal about patent value - and thus, on a more general level, by opaqueness of the patent system - if competing parties have an incentive to oppose a high-value patent *regardless* of whether it has been accelerated. Our model reveals that competing parties have an incentive to oppose a high-value patent regardless of its acceleration status as long as the probability that the patent is found valid in case of opposition is smaller than the cut-off value p_2^B. This value is determined by the relationship between the value of a non-accelerated high-value patent and the costs of patent opposition. The intuition here is that if the costs of opposition are sufficiently small in comparison to the value of a non-accelerated patent (which corresponds to how much competing parties get hurt in case the patent is granted), then competing parties have an incentive to oppose both accelerated and non-accelerated high-value patents.

Put together: If the probability p that a patent is found valid in case of opposition is not "too high" (that is, smaller than p_2^B), then changes in the behavior of both the patent applicant and rivals can be attributed to the omission of a signal about patent value. In this case, changes in the acceleration or opposition frequency due to the 2001 change in the EPO's information policy are evidence for opaqueness of the European patent system. Below we demonstrate that for our application we indeed expect p to be smaller than p_2^B.

■ **Predictions in case the patent system is opaque.** We start with the initial assumption that the European patent system is opaque with respect to patent value. The graphs in figures III.3 and III.4 show that in case the patent system is opaque and information about applicants' acceleration decisions gets concealed, depending on the underlying parameter relationships the behavior of applicants and rivals can change in different ways. The pattern of expected changes is as follows: In case the patent is either very weak ($0 < p < p_1^B$) or very strong ($p_3^B < p < 1$), we do not expect to observe changes in acceleration and opposition frequencies. In case the patent is of intermediate strength ($p_1^B < p < p_3^B$ - that is, for subsets I to IV of our parameter space), however, our model predicts both the frequency of patent applications and the frequency of opposition processes to change:

For small gains from acceleration of a high-value patent and small patent strength, we expect to observe only the acceleration frequency to increase (subset I). In case the probability that a patent is of high value is large, the increase in acceleration frequency is accompanied by an increase in opposition frequency (subset II). For large gains from acceleration and strong patents, we no longer expect to observe changes in acceleration frequency but only in opposition frequency. In case the probability that a patent is of high value is low, we expect the opposition frequency to decrease (subset III), whereas in case the probability of a high-value patent is high, we expect it to increase (subset IV). The reasons why these patterns develop were given in section III.3. Figure III.5 summarizes the results.

In order to derive specific predictions about the way acceleration and opposition frequencies change in reaction to concealment of the acceleration signal we need priors on our model parameters. These parameters are the values π_l and $\pi_h^{\neg a}$ of the patents, the probability θ with which a patent is of high value, the probability p with which a patent withstands opposition, the costs c_o and c_a of opposition respectively acceleration, and finally the value π_h^a of an accelerated high-value patent. In order to come up with sensible priors, in the following we shortly extract some stylized facts from the patent literature:

Regarding opposition costs c_o the literature is quite clear: According to for example Graham et al. (2002), who interviewed senior representatives of the European Patent Office, opposition costs c_o can be expected to be of a size of up to €0.1 m. With respect to the value of patents the literature is more ambiguous: The common

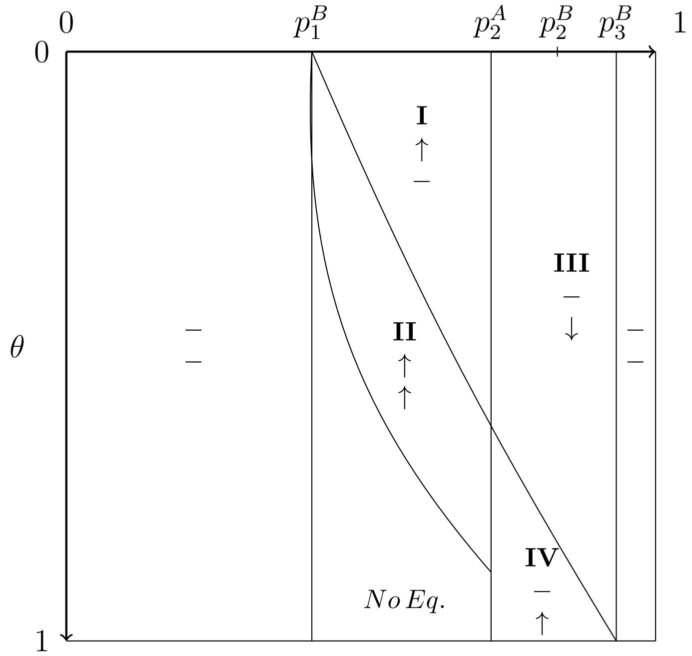

Figure III.5: **Expected changes in acceleration and opposition frequencies.** The figure shows the expected changes in acceleration and opposition frequencies in case the patent system is opaque with regard to patent value and information about applicants' acceleration requests gets concealed. The upper entry in each subset depicts the change in the frequency of acceleration requests, the lower entry the change in the frequency of oppositions. Depicted is the subset-partitioning for medium gains from acceleration. For high gains from acceleration subsets I and II disappear and only subsets III and IV remain between p_1^B and p_3^B.

finding here is that the distribution of patent value is heavily skewed - that is, the bulk of patents is of relatively low value, whereas a few patents are of quite high value. In studies on patent value the skewness of the value distribution is expressed in the fact that the median of the value distribution is usually found to be far smaller than its mean. However, due to different methodologies and data sets estimates of these two quantities range from magnitudes of below €0.1 m to an estimated median of €0.3 m and an estimated mean of €3 m in Gambardella et al. (2008). At the bottom line, the picture which emerges from studies on patent value is that the value of the bulk of patents seems to be close to the costs of opposition, whereas the value of a minority of patents exceeds the costs of opposition by more than one order of magnitude.

With respect to oppositions of granted patents, a study by Harhoff and Reitzig (2004) shows that the chances of successful opposition are good. In their sample, opposed patents were revoked in around one third and amended in 40% of all cases. Only in one fifth of the cases opposition was rejected. (In 10% of the cases the opposition procedure was closed due to unspecified reasons.) Note that the amendment of a patent can involve a narrowing of its scope, which might be counted as a (partial) success of the competing party. With respect to the procedure of accelerated patent examination information is scarce. Costs of acceleration costs solely arise due to the need to cooperate closely with the EPO in case accelerated examination is requested (there is no fee for accelerated examination), and thus should be quite small. There are no numbers on the additional profits an applicant can gain in case he requested accelerated examination and got his patent granted faster. However, the reduction in examination time can be substantial (from an average of around 40 month down to around 12 month), and thus the gains from accelerated examination should be economically significant. Put together, our stylized facts are:

SF1 While most patents are of low value, a minority of patents is of rather high value.

SF2 The value of low-value patents is in the range of the costs of opposition.

SF3 There is a good chance that opposition is successful.

SF4 The costs of accelerated patent examination are low, the gains substantial.

In order to derive predictions we operationalize these stylized facts by the following parameter assumptions: c_o=€ 0.1 m, π_l=€ 0.15 m, $\pi_h^{\neg a}$= € 1 m, c_a=€ 0.05 m and π_h^a=€ 2 m. Also, we expect θ to be smaller than 20% and p to be around 50%. It is important to note that with respect to the mechanics of our model the exact numerical values of the single parameters are not critical. What counts is the relationship between different parameters as expressed in our stylized facts.[7]

Figure III.6 depicts the parameter space which follows from these assumptions to scale. The gray area in figure III.6 marks the region where the probability θ that a given patent is of high value is between 0% and 20% and the probability p that a patent withstands opposition is around 50%. We do not mark a single point but use the fading grey area in order to symbolize that in our application we do not observe a single patent application process but many processes with different applicants and rivals involved. We expect the values of our model parameters to be different for each of these processes, but we make the assumption that the parameter

[7]In particular, these relationships are that between c_o and π_l, that between c_o and $\pi_h^{\neg a}$, and that between $\pi_h^{\neg a}$ and π_h^a. These relationships determine the relative positions of the cutoff-values p_1^B, p_2^B, p_3^B and p_2^A.

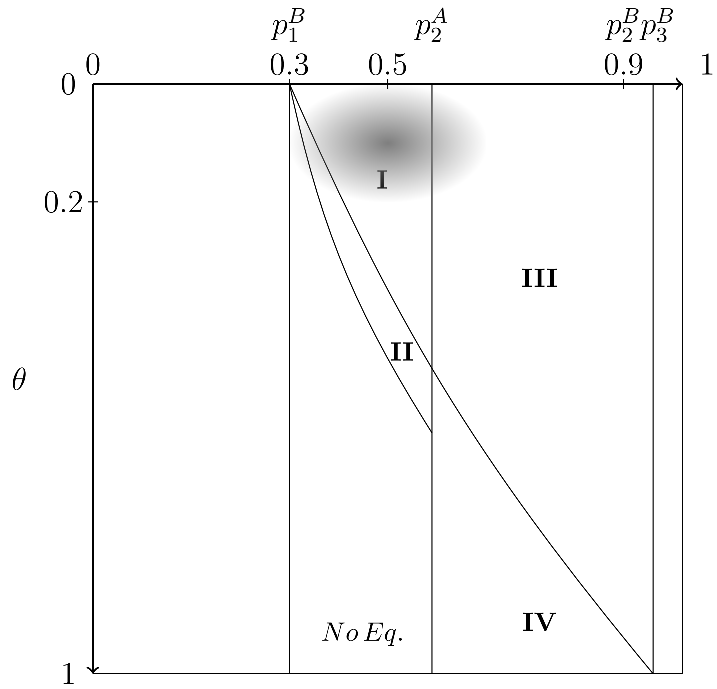

Figure III.6: **Parameter space (to scale) for parameter assumptions which mirror the stylized facts found in the literature.** The graph displays the parameter space for the parameter assumptions we derived from stylized facts about the patent application and opposition process in the European patent system. The parameter assumptions are c_o=€ 0.1 m, π_l=€ 0.15 m, $\pi_h^{\neg a}$= € 1 m, c_a=€ 0.05 m and π_h^a=€ 2 m, θ=20% and p=50%. As for each of the application and opposition processes in our data we expect the specific parameter values to be different but *on average* equal to our parameter assumptions, the fading grey area symbolizes where we expect most of the application and opposition processes to lie.

values of the different application processes do not vary strongly around the values we explicitly assumed. That is, the parameter values we explicitly assumed can be interpreted as the "average" parameter values in our application. Respectively, the fading grey area can be thought of to be a cloud of dots where each dot represents one particular application process.

We are interested in how the EPO's 2001 decision to conceal information about applicants' acceleration requests affected the behavior of applicants and rivals. In case the European patent system is opaque, we expect to observe the behavior of applicants and rivals to change. In particular, we expect the following changes:

H1 We expect a significant increase in the frequency of acceleration requests.

H2 We expect a significant decrease in the frequency of oppositions.

Year	Filings (#)	Accelerated search (%)	Accelerated examination (%)	Opposition (%)
1997	49,868	2.1	6.7	5.3
1998	53,350	2.6	7.4	5.1
1999	55,605	2.9	7.4	5.4
2000	59,193	3.0	7.0	5.1
2001	59,070	3.2	7.2	5.3
2002	55,822	3.2	7.3	4.9
2003	53,889	3.4	8.0	5.0
2004	51,323	3.7	8.8	4.7
2005	48,318	4.5	9.2	4.5
2006	44,321	5.0	9.5	4.3

Table III.2: **Yearly data on the number of filings and acceleration and opposition frequencies.** For each of the years 1997 to 2006, the table displays the number of filed patent applications (which were granted in the end), the fraction of these for which accelerated search was requested, the fraction for which accelerated examination was requested, and the fraction which was opposed after getting granted.

In the next subsection we use data on acceleration and opposition frequencies to put these hypotheses to the test.

Data and empirical results

The European Patent Office provided us with data on acceleration and opposition frequencies for the years 1997 to 2006. In particular, for each year we have information on the fraction of that year's filings for which applicants requested accelerated search, the fraction for which applicants requested accelerated examination, and the fraction which was opposed by rivals after getting granted. As our model abstracts both from applicants' decisions to withdraw their applications and the EPO's grant decision, and as we also do not have information about withdrawals in our data, we focus our analysis on filings which actually got granted later on. Each year, around 53,000 patent applications were filed and granted later on. For the years 1997 to 2006, table III.2 shows the number of applications filed, the fractions for which accelerated search and accelerated examination were requested, and the fraction of filings which were opposed after getting granted.

Table III.2 shows that the frequencies of accelerated search and accelerated examination exhibit a similar pattern over time. That is not surprising, as accelerated search is closely connected to accelerated examination: When a request for accel-

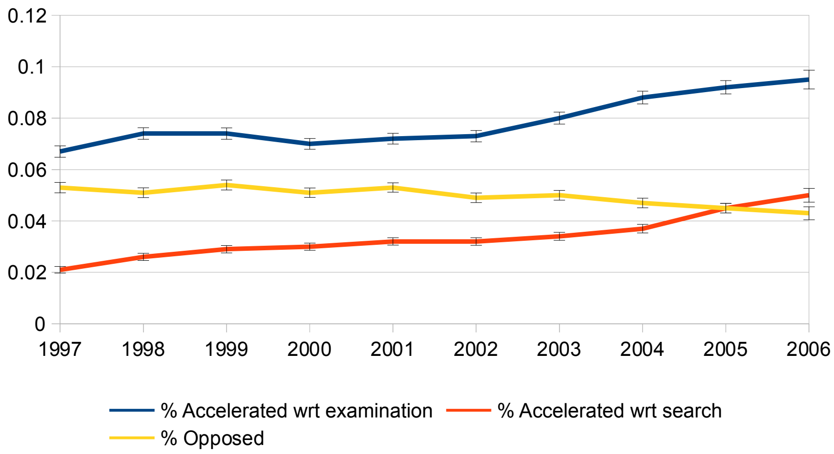

Figure III.7: **Frequencies of acceleration and opposition over time.** For the years 1997 to 2006, the graphs depict the fractions of patents for which accelerated examination were requested, the fractions for which accelerated search were requested, and the fractions which were opposed after getting granted. At each data point a 95% confidence intervals is displayed.

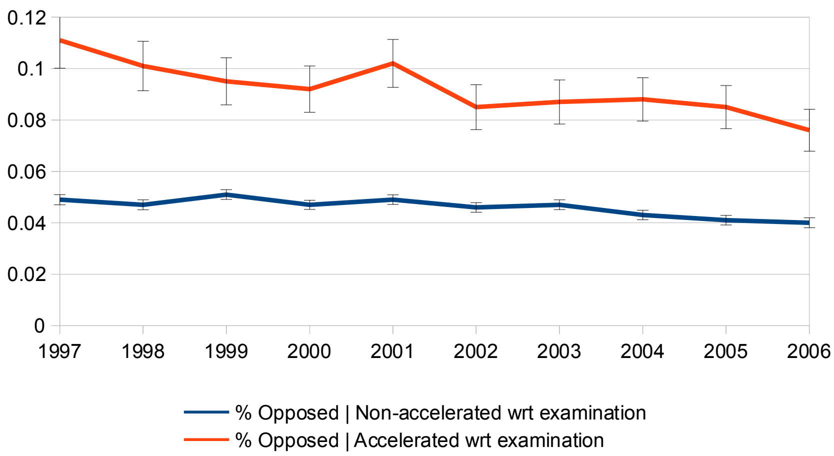

Figure III.8: **Frequency of opposition conditional on acceleration status.** For the years 1997 to 2006, the graphs display the development of opposition frequencies over time for the fractions of patents for which accelerated examination were requested and for the fractions for which there were no request. At each data point a 95% confidence intervals is displayed.

erated examination is filed together with a patent application, this automatically induces an accelerated search process. For the sake of clarity, in the following we will display data for both the frequencies of accelerated search and accelerated examination, but focus our discussion mainly on the frequency of requests for accelerated examination.

The graphs in figure III.7 show the developments of acceleration and opposition frequencies over time. Up to the EPO's policy change in December 2001 there seems to be no significant change in acceleration and opposition frequencies. After the EPO's decision to conceal information about applicants' acceleration requests, however, there is a clear increase in the frequency of accelerated examination and a little less pronounced decrease in opposition frequency. Figure III.8 takes a closer look at the development of the opposition frequency over time. There, the graphs display the development of opposition frequencies for the fraction of patents for which accelerated examination was requested and for the fraction of patents for which there was no request. Whereas there seems to be no change in opposition frequency for the fraction of patents which were not accelerated, the opposition frequency for the fraction of accelerated patents dropped considerably after the EPO's 2001 policy change.

In order to check whether the observed changes in acceleration and opposition frequencies are significant, we interpret the EPO's 2001 policy change as "treatment" and divide our data into a "pre-treatment" period covering the years 1997 to 2000 and a "post-treatment" period covering the years 2002 to 2006. We leave out the year 2001 because the EPO announced to change its information policy in October 2001 and we have only yearly data available. Thus, we do not know which fraction of the 2001 filings was affected by the EPO's policy change. In the first line of table III.3 we report the p-values for a two-sample t-test. It tests the hypothesis that the frequency in the pre-treatment period is equal to the frequency in the post-treatment period. We do this test for the frequencies of accelerated examination requests and oppositions. With respect to oppositions we do the pre-post-comparison for three frequencies: The unconditional frequency of opposition, the frequency conditional on accelerated examination had been requested, and the frequency conditional on there had been no request for accelerated examination. For all acceleration and opposition frequencies, it shows that all differences are statistically significant (with respect to at least a 0.1% significance niveau). In addition, the differences in the frequencies of requests for accelerated examination and the difference in the frequency of oppositions conditional on accelerated examination had been requested are not only of statistical but also of economic significance - the differences are in the order of magnitude of around one percentage point. This is large compared to the level of the frequencies, which ranges from around two to ten percent. Whereas being statistically significant, the changes in the frequencies of opposition unconditional

	H_0: Freq. pre-treatment period = Freq. post-treatment period.			
	Accelerated examination	Opposition	Opposition (not accelerated)	Opposition (accelerated)
1997-00 vs. 2002-06	0.071 vs. 0.085 <0.001	0.052 vs. 0.047 <0.001	0.048 vs. 0.044 <0.001	0.100 vs. 0.084 <0.001
2000 vs. 2002	0.070 vs. 0.073 0.024	0.051 vs. 0.049 0.298	0.047 vs. 0.046 0.439	0.092 vs. 0.085 0.273
1999 vs. 2002	0.074 vs. 0.073 0.559	0.054 vs. 0.049 <0.001	0.051 vs. 0.046 <0.001	0.095 vs. 0.085 0.115
1998-99 vs. 2002-03	0.074 vs. 0.076 0.026	0.053 vs. 0.050 0.001	0.049 vs. 0.046 0.007	0.098 vs. 0.086 0.009

Table III.3: **Statistical significance of the changes in acceleration and opposition frequencies.** The table displays the p-values of two-sample t-tests. The t-tests were performed with respect to the frequencies of requests for accelerated examination and oppositions (both unconditional and conditional on the acceleration status with respect to examination). For each test, the null hypothesis is that frequencies are equal for the respective pre- and post-treatment periods.

on the acceleration status and conditional on that there had been no request for accelerated examination are of minor economic importance.

The fact that the differences in frequencies are statistically significant if we define the pre-treatment period to cover the years 1997 to 2000 and the post-treatment period to cover the years 2002 to 2006 is actually not surprising: Because we cover a long period of time, the number of observations and thus estimation efficiency is very high. This gain in efficiency due to a long observation period, however, comes with a major drawback: The more years we cover, the higher is the possibility that we capture events in the development of the European patent system which are unrelated to the EPO's 2001 policy change, but which might have had an effect on acceleration and opposition frequencies. That is, the longer the time period, the higher the possibility that there is some bias in our results. In order to tackle this potential problem we test whether there remains a statistically significant change in frequencies if we define shorter pre- and post-treatment periods.

When we compare only frequencies of the years 2000 and 2002, we see that the changes in opposition frequencies become insignificant. This might be due to the loss of observations and thus estimation power. However, we suspect the reason is different and indeed structural: The EPO recommends applicants to request accelerated examination either when they file the patent application or when they receive the search report. On average, applicants receive the search report 18 month after the filing of their application. That means, if an applicant filed an application in 2000, it might well be that he received his search report after December 2001. In this case, with respect to his acceleration decision the applicant would had been affected by the EPO's 2001 policy change. That is, filings of the year 2000 might well have been affected by the EPO's 2001 policy change, and in this case we would expect to observe no changes in frequencies. Thus, instead of the years 2000 and 2002 we compare the years 1999 and 2002. Now, we indeed see that changes in frequencies are significant (on a 5% significance niveau). The only exemption is the change in the frequency of requests for accelerated examination. However, with a look at figure III.7, we see that there is an irregular drop in the frequency of accelerated examination in the year 2000. Also, acceleration frequencies begin to significantly rise only from the year 2003 on.[8] If we include one more year and compare the years 1998-1999 to 2002-2003, also the change in the frequency of requests for accelerated examination becomes significant again.

In summary, between the years before and those after the EPO's 2001 policy change we observe the frequency of requests for accelerated examination to increase. The increase is significant both statistically and economically. Although statistically significant, we observe only a slight decrease of the frequency of oppositions unconditional on acceleration status. If we condition the frequency of oppositions on whether accelerated examination had been requested, we still observe only a slight decrease of the frequency of oppositions against patents for which there had been no requests for accelerated examination. However, for patents for which accelerated examination had been requested our data shows an economically significant drop in the frequency of oppositions.

Discussion and Interpretation

Our model predicts acceleration and opposition frequencies to change when information about applicants' acceleration decisions gets concealed. In our data we indeed observe that acceleration and opposition frequencies are significantly different between the years before the EPO's policy change in 2001 and the years afterwards. However, a necessary condition to interpret these changes in the light of our model

[8] This might be related to the facts that applicants became aware of the EPO's policy change only after the EPO published a President's Notice in October 2001 and that the EPO advices to file requests for accelerated examination either when submitting the patent application or after the receipt of the search report (which happens around 18 month after the initial filing of a patent application).

is that these changes are actually *caused* by the 2001 change in EPO's information policy. Thus, before we turn to an interpretation of these changes in the light of our model, we first shortly discuss under what conditions the observed changes in acceleration and opposition frequencies can be causally linked to the EPO's 2001 policy change, and whether these conditions are likely to be met in our data.

■ **Causality.** In principle, we could establish a causal link between changes in behavior over our observation period and the EPO's policy change in 2001 if our data stemmed from an ideal experiment where both applicants and rivals were randomly assigned to either a patent regime with disclosed acceleration information (the non-treated control group) or a patent regime with concealed acceleration information (the treatment group). Given large enough numbers, this setup would ensure two things: First, the composition of the non-treated and the treated group would on average be the same. Second, applicants respectively rivals would not have the possibility to select themselves into treatment. That is, given this ideal setup, if hypothetically both the control and the treatment group were *not* treated, acceleration and opposition frequencies would on average be the same for both groups. This in turn ensures that if we observed differences between the treatment and the control group, we could interpret these as being causal effects of the treatment (that is, in our case, of concealment of information about applicants' acceleration requests).

The situation in our data differs from this ideal setup: Our data does not stem from an ideal experiment but is quasi-experimental. Instead via a random generator which by chance assigns applicants and rivals to either the treatment or the control group, treatment (that is, the EPO's 2001 decision to conceal the acceleration signal) happened in the time dimension and affected all parties without exception. Thus, the control group is comprised of all applicants and rivals active before the EPO's new information policy came into effect in 2001, and the treatment group is comprised of all applicants and rivals active afterwards. In order to establish causality we have to check two conditions: The composition of applicants and rivals before the 2001 change in EPO's information policy has to be the same as that afterwards (on average). Also, there must be no external shock unrelated to the treatment which affected our outcomes of interest (that is, applicants' propensity to accelerate and rivals' propensity to oppose).

So far the EPO has provided us with data on a quite aggregate level. This data does not inform us about the identity of applicants and rivals, and thus it does not allow us to explicitly control for changes in the composition of firms before and after the EPO's 2001 policy change. However, judging from the EPO's annual reports, no major structural break in the European patent system seems to have occurred during the years 1997 to 2006.[9] As our observation period is also rather short, we thus argue that it is not unreasonable to assume that the composition of applicants and rivals before the 2001 change in the EPO's information policy is not significantly different from that afterwards. In addition, the graphs in figure III.8 show that there was no economically significant change in the frequency of oppositions against patents which had not been accelerated. As the frequency of oppositions against patents which had not been accelerated should not be affected by the EPO's 2001 policy change, this is an indication that there also were no structural breaks in the behavior of applicants and rivals during our observation period.

Put together, we argue that for our quasi-experimental setting it holds that both the composition of applicants and rivals before the 2001 change in EPO's information policy is not significantly different from the composition afterwards, and that apart from the EPO's 2001 policy change there were no other shocks which systematically influenced firms' propensities to accelerate respectively oppose patents. Thus, we conclude from the data from our quasi-experiment that there is a causal link between the EPO's 2001 decision to conceal information about applicants' acceleration requests and the observed changes in acceleration and opposition frequencies.

■ **Evidence for opaqueness of the European patent system.** Our model predicts that in case the European patent system is opaque with respect to the value of patents, in reaction to the EPO's 2001 policy change the frequency of requests for accelerated examination should increase and the frequency of oppositions should decrease. These were hypotheses one and two we derived above. Our data confirms the predictions of our model: In reaction to the EPO's 2001 policy change we indeed observe the acceleration frequency to increase and the opposition frequency to decrease. This supports our assumption that the European patent system is opaque with respect to patent value. That is, without a signal transporting information

[9]For example, during our observation period the distribution of applications over the ten most frequent residence countries did not significantly change. Also, there was no major change in the distribution of applications over the ten most frequent technical fields around the year 2001: Apart from a slight increase in the fraction of applications from the field of information technology, the distribution of applications over the technical fields remains nearly unchanged. A breakdown of applications by residency and technical field is given in appendix C.3.

about patent value rivals seem to have difficulties to identify a patent's contribution merely from the conventional data generated by the patent office.[10]

□ **Welfare implications of opaqueness.** According to our model, the welfare implications of the fact that the European patent system is opaque with respect to the value of patents depend on the distribution of high- and low-value patents. Above, in order to derive hypotheses H1 and H2, we assumed that the majority of patents is of low-value. As our data fits the predictions of our model, our results support this assumption. Proposition 3 states that in case the probability that a patent is of high value is low, aggregate welfare of applicants and rivals is (weakly) higher in case the patent system is opaque than in case the patent system is transparent. The reason for this somewhat counterintuitive result is that in case the occurrence of high-value patents is unlikely and the patent system is opaque, applicants can hide high-value patents among the bulk of low-value patents. As discussed in section III.3, from an aggregate welfare perspective this saves either acceleration costs (in case patents are weak) or opposition costs (in case patents are strong). We conclude that for the case of the European patent system opaqueness with respect to patent value might not be detrimental to the aggregate welfare of applicants and rivals. However, as our model focuses on the interplay between applicants and rivals during the patent application process, this of course is a partial welfare result only. It abstracts completely from the implications of opaqueness of the patent system on the innovation process in society.

□ **Welfare implications of concealment of acceleration information.** We established that the European patent system is opaque with respect to patent value. In section III.3, we discussed that for the case of an opaque patent system the implications of the EPO's 2001 decision to conceal acceleration information depend on the values of the underlying structural parameters. The patterns we observe in the data support our assumption that in the European patent system the occurrence of a high-value patent is rather unlikely. For low probabilities of high-value patents the implications of concealment of acceleration information on the aggregate welfare of applicants and rivals are ambiguous: In case patents are weak, our model predicts aggregate welfare to decrease, whereas in case patents are strong, it predicts aggregate welfare to increase.

[10] However, note that the graphs in figure III.8 show that after the EPO concealed information about acceleration requests in 2001, we still observe accelerated patents to be opposed more frequently than non-accelerated patents. If we entertain the assumption that acceleration is only worthwhile for high-value patents, this indicates that the European patent system transmits some information about the value of patents, but that the transmission mechanism is imperfect.

As discussed in more detail in section III.3, the reason for this ambiguity is that in case of weak patents concealment of acceleration information allows applicants to accelerate their high-value patents in disguise, which, from an aggregate welfare perspective, produces unnecessary acceleration costs. In case of strong patents, concealment of acceleration information no longer allows rivals to target high-value patents. Thus, rivals refrain from opposition, and this saves opposition costs (from an aggregate welfare perspective).

Figure III.6 shows that the boundary line which separates weak from strong patents (that is, subset I from subset III) runs through the area of the parameter space where we expect most of the application processes in our data to lie. (This area is marked by the grey shading in figure III.6.) As just stated, in case acceleration information gets concealed, for weak patents we expect aggregate welfare to decrease, whereas for strong patents we expect it to increase. Thus, we cannot give a definite answer to the question how the EPO's 2001 decision to conceal acceleration information affected aggregate welfare of applicants and rivals. The answer to this question is determined by the location of the boundary line between subset I and subset III relative to the mass of application processes. This location depends on how much the value of a patent increases in case it is accelerated. Unfortunately, the value of acceleration is the one model parameter about which information is very scarce in the literature. Therefore, we are quite uncertain with respect to the exact location of the boundary line. To produce figure III.6, we made the assumption that acceleration doubles the value of a high-value patent. If in fact the value of a patent increases by less than the factor two when it is accelerated, then the position of the boundary line would be further to the right, and we would conclude that the EPO's 2001 policy change decreased aggregate welfare. In contrast, if acceleration increased the value of a patent by more than the factor two, the boundary line would move to the left, and we would conclude that the EPO's 2001 policy change increased aggregate welfare. That is, in order to give a reliable answer to the question how the EPO's 2001 decision to conceal information about acceleration requests affected the aggregate welfare of applicants and rivals, we need to collect more information about the value of patent acceleration.

III.5 Conclusion

This chapter adds to a better understanding of the fundamental deal between the applicant of a patent and society - that is, the granting of exclusion rights in exchange to disclosure of technical knowledge - by asking whether the European patent system is indeed transparent with respect to the value of a patented innovation or instead rather opaque. We try to give an answer to this question by exploiting a rule change in the European patent system: While before December 2001 applicants' requests for accelerated search and examination were disclosed to the public, afterwards these requests were treated as confidential information. We developed a model of the patent application and opposition process which shows that concealment of the acceleration signal leads to specific changes in the behavior of applicants and rivals in case the European patent system is opaque with respect to patent value. In particular, in reaction to the EPO's 2001 policy change, the frequency of acceleration requests should increase and that of oppositions should decrease. These predictions of our model are met by the data, which gives support to our presumption that the European patent system is opaque with respect to the value of patents. That is, the main conclusion we draw from our analysis is that it seems difficult to identify a patent's contribution solely on the basis of the conventional data generated by the EPO.

Based on our main finding that the European patent system is opaque with respect to patent value, we then took first steps towards an assessment of the welfare implications of opaqueness and of the availability of a signal transporting value information. Perhaps surprisingly we find that a transparent patent system might not always be preferable to an opaque one. In fact, in case of the European patent system, opaqueness with respect to patent value seems to increase the combined welfare of applicants and rivals, both for the case where there is a publicly observable acceleration signal and the case where there is not. However, this is only a partial welfare result, as our analysis focuses on the firms directly involved in the patent application process and does not take into account the effects opaqueness on the progress of innovation in society. Also, our theoretical analysis focuses on opaqueness of the patent system with respect to the value of a given patent. If there is another dimension of opaqueness of the patent system in the sense that firms have to make considerable investments in order to identify potentially conflicting prior art - that is, if there are substantial search costs, then a welfare evaluation of opaqueness solely based on the model presented in this chapter might be too limited in its scope.

The welfare implications of the availability of a signal transporting value information (given an opaque patent system) depend critically on how strongly a patent applicant's profits increase in case of accelerated patent examination. For the case of the European patent system, it seems that the stronger the increase in profits from acceleration, the more likely it is that the EPO's 2001 change in its information policy increased aggregate welfare. In order to come up with a more substantiated statement regarding the welfare implications of the EPO's 2001 policy change, however, we need reliable estimates of the value of patent acceleration. Unfortunately, to the best of our knowledge the issue of patent acceleration has not received much attention in the literature yet. Thus, in order to assess the full implications of opaqueness and the availability of a value signal, more research, especially on the empirics of patent acceleration, is called for.

Appendix A

Umbrella Branding and Consumer Inertia

A.1 Details on the prior specification

I specify a two-stage prior as in Rossi et al. (2005) and Dube et al. (2010). The reason for specifying a two-stage prior is that a two-stage prior is flexible in the sense that it accounts for heterogeneity in the household-level parameters $\boldsymbol{\theta}_h$. Or, in other terms: It a priori assumes households to be different and thus allows the posterior to accommodate household heterogeneity far more efficient than a simple one-stage prior.

In particular, I assume my model coefficients $\boldsymbol{\theta}_h = (\boldsymbol{\alpha}^h, \eta^h, \beta^h, \gamma^h)$ to be household-specific and to follow a mixture of normals distribution:

$$p(\boldsymbol{\theta}_h|\boldsymbol{\pi}, \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\}) = \sum_{k=1}^{K} \pi_k \phi(\boldsymbol{\theta}_h|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) \tag{A.1}$$

Equation (A.1) represents the so-called first-stage prior for my model coefficients $\boldsymbol{\theta}_h$: I assume my model coefficients to follow a mixture of K multivariate normal distributions, each with mean $\boldsymbol{\mu}_k$ and variance matrix $\boldsymbol{\Sigma}_k$. Each multivariate normal distribution is weighted with π_k.

A so-called second-stage prior further specifies the components of the first-stage prior. In particular, it specifies the number of components K, and that the π_k, the $\boldsymbol{\mu}_k$ and the $\boldsymbol{\Sigma}_k$ are drawn from certain distributions with certain parameters. These distributions are chosen in a way which simplifies computations in the iterations of the Markov Chain Monte Carlo algorithm (in essence, these distributions are easy to simulate from):

The π_k shall be drawn from a multivariate Dirichlet distribution with concentration parameter $a > 0$. The concentration parameter determines how on average the K components are mixed. A concentration parameter above one has as consequence that on average the mixture of normals is a balanced composition of all components, a concentration parameter below one has as consequence that on average the mixture of normals is dominated by a certain component.

$\mathbf{\Sigma}_k$, the variance matrices of the multivariate normal components, shall be drawn from an inverse Wishart distribution with scale matrix $\mathbf{V}$ and ν degrees of freedom. It has to hold that $\nu > p - 1$, where p equals the dimensions of $\mathbf{\Sigma}_k$ and $\mathbf{V}$. The expected value of $\mathbf{\Sigma}_k$ is

$$\mathrm{E}[\mathbf{\Sigma}_k] = \frac{\mathbf{V}}{\nu - p - 1}. \tag{A.2}$$

$\boldsymbol{\mu}_k$, the means of the multivariate normal components, shall be drawn from a normal distribution with mean $\bar{\boldsymbol{\mu}}$ and variance matrix $\mathbf{\Sigma}_k \mathbf{A}_\mu^{-1}$. $\mathbf{A}_\mu$ is a matrix which allows to scale the precision of the prior on the means of the multivariate normal components: In general, the larger the entries of $\mathbf{A}_\mu$, the higher the precision of the prior.

As prior values I choose $K = 5$, $\nu = 29$, $\mathbf{V} = 10\mathbf{I}$, $\mathbf{A}_\mu = \frac{1}{16}\mathbf{I}$, $\bar{\boldsymbol{\mu}} = \mathbf{0}$ and $a = 0.1$. $\mathbf{I}$ stands for the unit matrix, $\mathbf{0}$ for the zero vector. With these my prior specifications can concisely be written as

$$\begin{aligned}
\text{1st stage:} \quad & p(\boldsymbol{\theta}_h|\boldsymbol{\pi}, \{\boldsymbol{\mu}_k, \mathbf{\Sigma}_k\}) = \sum_{k=1}^{K} \pi_k \phi(\boldsymbol{\theta}^h|\boldsymbol{\mu}_k, \mathbf{\Sigma}_k) \\
\text{2nd stage:} \quad & \boldsymbol{\pi} \sim \text{symm. Dirichlet}(a) \\
& \boldsymbol{\mu}_k|\mathbf{\Sigma}_k \sim N(\bar{\boldsymbol{\mu}}, \mathbf{\Sigma}_k \mathbf{A}_\mu^{-1}) \\
& \mathbf{\Sigma}_k \sim IW(\nu, \mathbf{V}) \\
& K = 5, a = 0.1, \mathbf{A}_\mu = 1/16\mathbf{I}, \\
& \bar{\boldsymbol{\mu}} = \mathbf{0}, \nu = 29, \mathbf{V} = 10\mathbf{I}
\end{aligned}$$

I choose the number of normal components to be 5 in order to being able to capture household heterogeneity without overfitting the data and without running into computational troubles. The second-stage priors for ν, $\mathbf{V}$, $\mathbf{A}_\mu$, $\bar{\boldsymbol{\mu}}$ and a are chosen in a way which renders the first-stage prior on $\boldsymbol{\theta}_h$ non-informative and diffuse in expectation (in expectation, the prior mean of $\boldsymbol{\theta}_h$ is zero, and the variance of each normal component of each coefficient is five). In section I.5 these prior specifications and the robustness of my estimation results against different possible prior specifications are discussed in more detail.

A.2 Details on MCMC algorithm

The posterior distribution of my model coefficients $\boldsymbol{\theta}_h$ is given as

$$p(\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_H | \mathbf{y}_1, ..., \mathbf{y}_H, \mathbf{h}) \sim \left[\prod_h p(\mathbf{y}_h | \boldsymbol{\theta}_h) p(\boldsymbol{\theta}_h | \boldsymbol{\tau}) \right] \cdot p(\boldsymbol{\tau} | \mathbf{h}), \tag{A.3}$$

where $\boldsymbol{\tau} = (\boldsymbol{\pi}, \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\})$ and $\mathbf{h} = (K, a, \mathbf{A}_{\boldsymbol{\mu}}, \bar{\boldsymbol{\mu}}, \nu, \mathbf{V})$. $\mathbf{y}_h$ denotes the data available for household h. This posterior is of a form for which analytical results regarding its moments and marginals are not available. Thus I have to retreat to the use of simulations to get insights into its features. The simulation method used is a Markov Chain Monte Carlo Method. The basic idea here is very simple: Starting from initial conditions $\boldsymbol{\Theta}^0 = (\boldsymbol{\theta}_1^0, ..., \boldsymbol{\theta}_H^0)$ the method uses certain rules to generate a sequence $\boldsymbol{\Theta}^0, \boldsymbol{\Theta}^1, \boldsymbol{\Theta}^2, ...$ of household coefficients. The sequence rules ensure that the long-run distribution of the $\boldsymbol{\Theta}^r$ converges against the posterior distribution $p(\boldsymbol{\Theta} | \mathbf{y}_1, ..., \mathbf{y}_H, \mathbf{h})$. As the long-run distribution of the $\boldsymbol{\Theta}^r$ converges against the posterior distribution $p(\boldsymbol{\Theta} | \mathbf{y}_1, ..., \mathbf{y}_H, \mathbf{h})$, I can use the sequence of the $\boldsymbol{\Theta}^r$ to simulate moments and marginals of the posterior.

The MCMC algorithm I use is a hybrid Markov chain algorithm. This algorithm is described in great technical detail in Rossi et al. (2005). The algorithm makes use of the fact that the prior on the household coefficients $\boldsymbol{\Theta}$ is defined in two stages: The distribution of the household coefficients $\boldsymbol{\Theta}$ depends on the parameters $\boldsymbol{\tau} = (\boldsymbol{\pi}, \{\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k\})$ of the mixture of multivariate normals distribution. Via the likelihood of the data the coefficients $\boldsymbol{\Theta}$ are furthermore linked to the household data $\mathbf{y} = (\mathbf{y}_1, ..., \mathbf{y}_H)$. On the second stage in turn the $\boldsymbol{\Theta}$ can be interpreted as data regarding the draw of household-level parameters. These conditional dependencies can be written as

$$\boldsymbol{\Theta} \mid \boldsymbol{\tau}, \mathbf{y} \tag{A.4}$$

$$\boldsymbol{\tau} \mid \boldsymbol{\Theta}. \tag{A.5}$$

The hybrid Markov chain algorithm proceeds in two steps: In the first step it uses relationship (A.4) to draw household coefficients $\mathbf{\Theta}$ using household data $\mathbf{y}$ and a former draw of $\boldsymbol{\tau}$. The draw itself is done by a Metropolis algorithm, which employs each households' multinomial logit likelihood and an accept/reject like draw method. In the second step it uses relationship (A.5) and the draw of $\mathbf{\Theta}$ to draw $\boldsymbol{\tau}$. This draw is done using a standard unconstrained Gibbs sampler. The second-stage prior specifications ensure that the conditional distribution (A.5) is known and Gibbs sampling is therefore possible. Starting from initial conditions, the hybrid Markov chain algorithm iterates through steps one and two. It can be shown that in the long-run the distribution of the draws generated by the algorithm converges against the posterior distribution (A.3). This allows me to use the long-run sequence of draws to estimate marginals and moments of the posterior.

The hybrid Markov chain algorithm is implemented in computer code which can be found in a contributed R package called `bayesm`. This package is described in Rossi et al. (2005) and has been developed by these authors. It can be found on the CRAN network of mirror sites (http://cran.r-project.org). The function the algorithm is implemented in is `rhierMnlRwMixture`. I modified this function to allow for varying choice sets. Concretely, I modified the functions `llmnl` and `mnlHess`, which are called by `rhierMnlRwMixture`, such that they are able to deal with a non-constant choice set. For each function, the modification is straightforward - instead of using fast matrix operations (which rely on a constant choice set) to compute the log-likelihood respectively the Hessian of the multinomial model I use loops over all units in order to be able to account for differing choice sets. The drawback with my modification is that it drastically increases computation time. I produced my results with the modified function. For the simulation of the posterior distributions I used $10,000$ iterations with a burn-in period of $1,000$ iterations.

A.3 Bayesian model comparison

Comparison of different models is straightforward in the Bayesian framework. As detailed in Rossi et al. (2005), model choice based on posterior model probabilities is in line with decision theory if there is a zero-one loss function (0 in case the true model is chosen, 1 otherwise). The posterior probability of a model M_i is given as

$$p(M_i|D) = \frac{p(D|M_i)p(M_i)}{p(D)}. \tag{A.6}$$

$p(M_i)$ denotes the prior probability of model M_i, $p(D)$ the (unconditional) probability of observing the data at hand, and $p(D|M_i)$ the marginal likelihood for M_i. $p(D)$ is independent of model M_i. If I assume equal prior probabilities for all the models I want to compare, then it holds that

$$p(M_i|D) \sim p(D|M_i). \tag{A.7}$$

That means that under the assumption of equal prior model probabilities model choice based on posterior model probabilities boils down to model choice based on marginal model likelihoods.

The marginal model likelihood is defined as

$$p(D|M_i) = \int p(D|\boldsymbol{\theta}, M_i)p(\boldsymbol{\theta}|M_i)\mathrm{d}\boldsymbol{\theta}. \tag{A.8}$$

As shown by Dube et al. (2010), using the method of Newton and Raftery (1994) the marginal model likelihood can be estimated as

$$\hat{p}(D|M_i) = \left(\frac{1}{R}\sum_{r=1}^{R}\frac{1}{p(D|\boldsymbol{\Theta}^r, M_i)}\right)^{-1}. \tag{A.9}$$

For the comparison of different models I use the logarithm of the marginal model likelihood. Note finally that posterior model probabilities automatically correct for different parameter dimensions.

A.4 Robustness checks: Posterior distributions

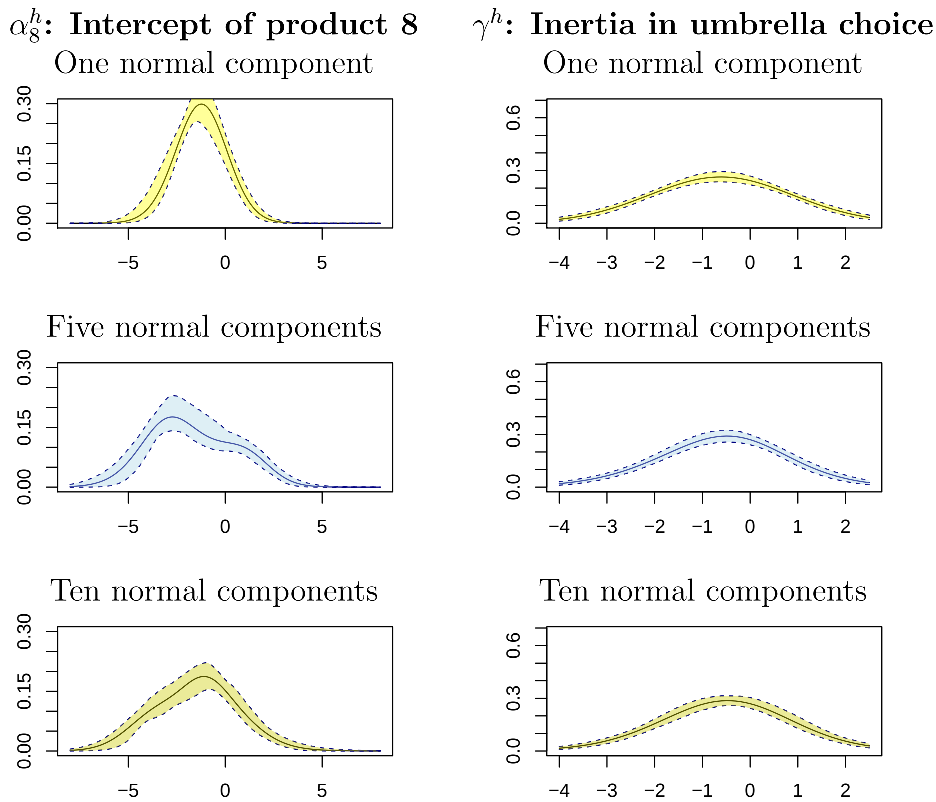

Figure A.1: **Comparison of posterior distributions derived with different numbers of normal components.** Each graph depicts the pointwise posterior mean and the 95% credibility region of marginal posterior densities. The left column depicts marginal posterior densities of the intercept of product eight, the right column depicts those of the coefficient γ^h, which captures household inertia in umbrella brand choice. The results were derived with different priors regarding the number of normal components: one, five and ten normal components. Apart from the normal components the prior specifications equal those detailed in appendix A.1. The results are based on 5,263 purchasing observations of 775 households.

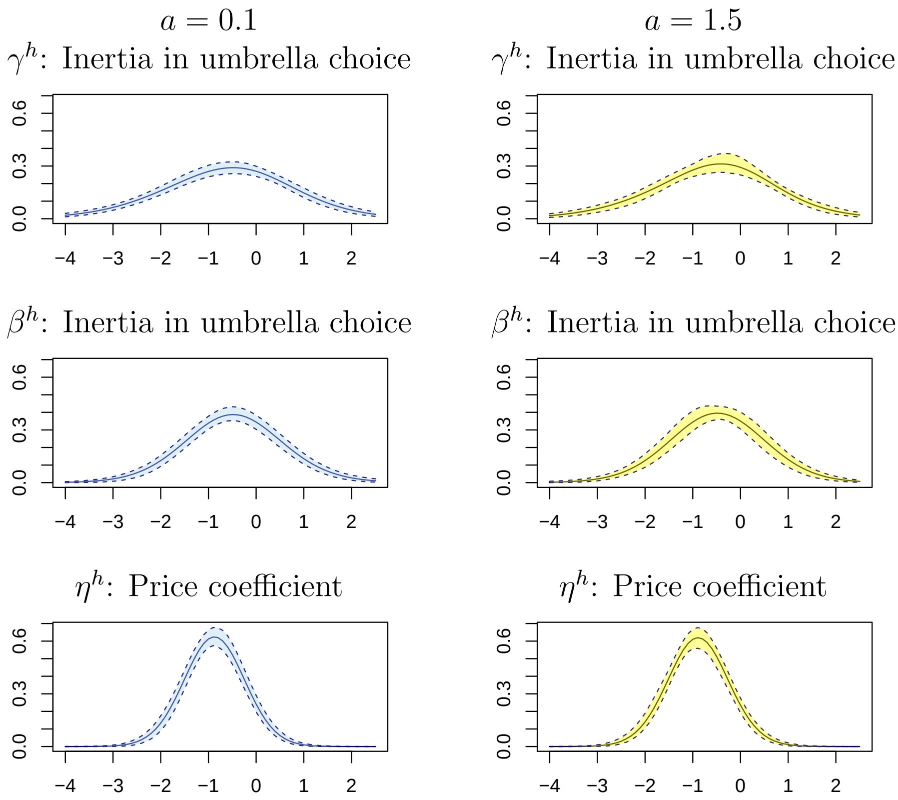

Figure A.2: **Comparison of posterior distributions derived with different prior specifications regarding the concentration parameter** a**.** All graphs depict pointwise posterior means and 95% credibility regions of marginal posterior densities. The graphs in the left column depict results derived with my base prior specification, which assumes the concentration parameter a to equal 0.1 My base prior specification is detailed in appendix A.1; these results are equal to those depicted in figure I.3. The graphs in the right column depict results derived with a prior specification which assumes the concentration parameter a to equal 1.5. The results are based on 5,263 purchasing observations of 775 households.

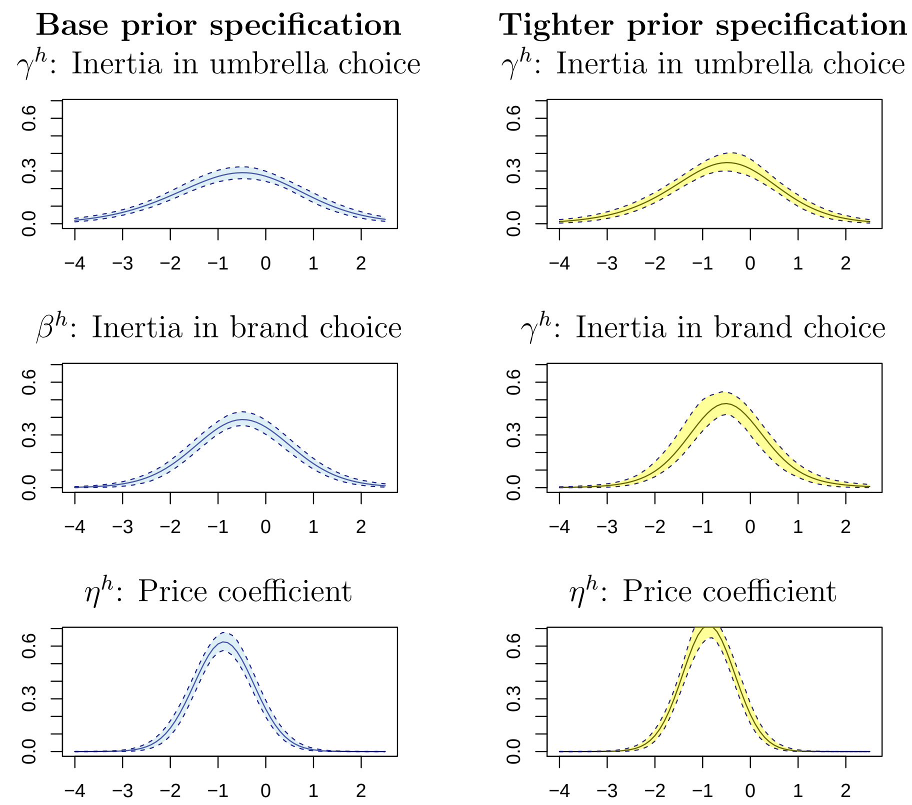

Figure A.3: **Comparison of posterior distributions derived with base prior specification to posterior distributions derived with tighter prior specification.** All graphs depict pointwise posterior means and 95% credibility regions of marginal posterior densities. The graphs in the left column depict results derived with my base prior specification. My base prior specification is detailed in appendix A.1; these results are equal to those depicted in figure I.3. The graphs in the right column depict results derived with a prior specification which assumes the coefficient variances to be in expectation a priori half as large as with the base prior specification. The results are based on 5,263 purchasing observations of 775 households.

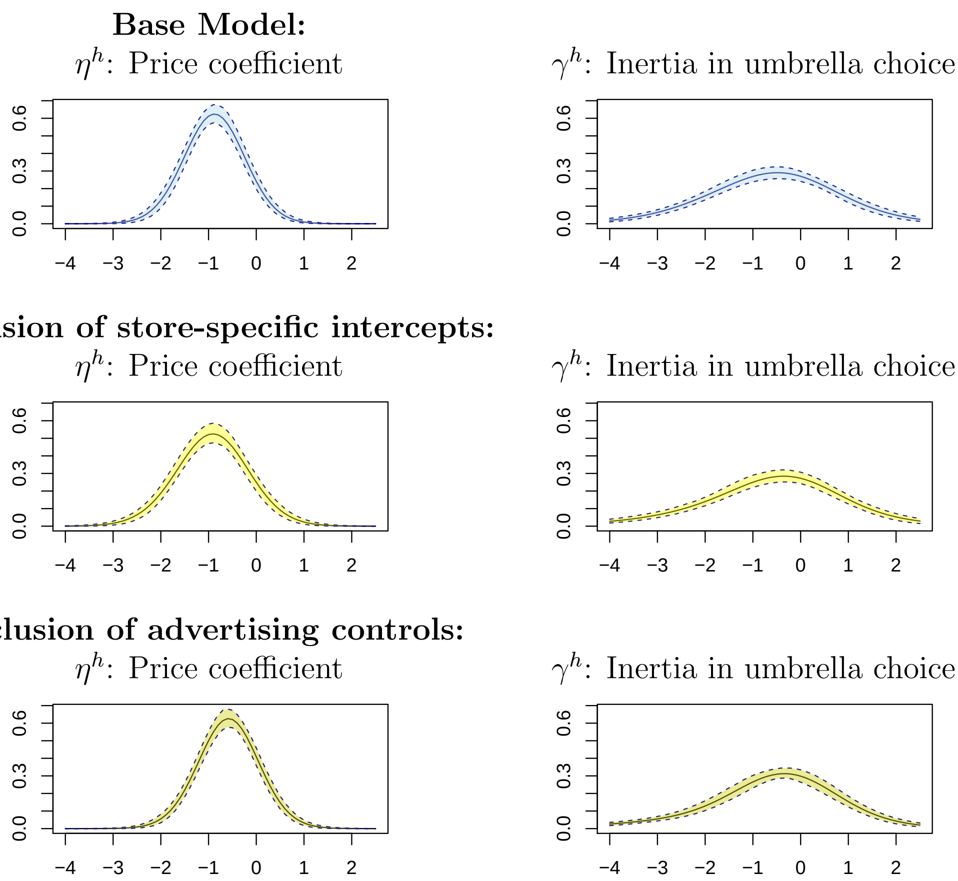

Figure A.4: **Comparison of posterior distributions derived with the base model, the base model extended by store-specific intercepts, and the base model extended by advertising controls.** All graphs depict pointwise posterior means and 95% credibility regions of marginal posterior densities. Depicted are the marginal posterior distributions of the price coefficient and the umbrella brand choice inertia coefficient. From top to bottom these posteriors were derived from the base model (I.1), the base model extended by store-specific intercepts, and the base model extended by advertising controls. All results were derived using the base prior specifications given in appendix A.1 and are based on 5,263 purchasing observations of 775 households.

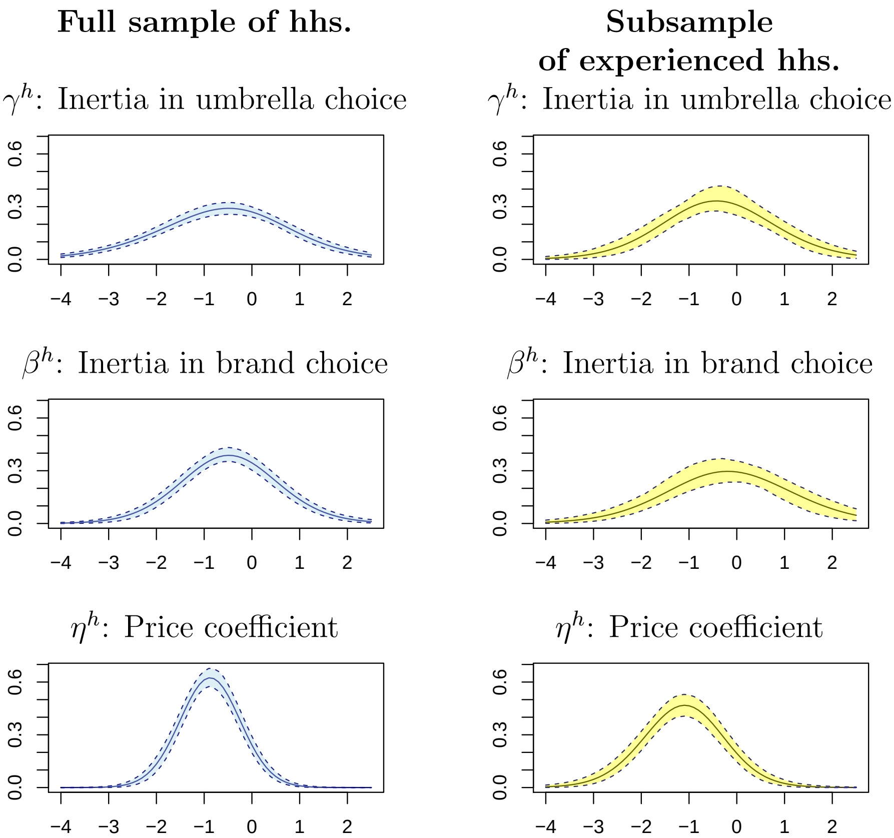

Figure A.5: **Comparison of posterior distributions derived with the full sample of households to posterior distributions derived with a selected sample of experienced households.** All graphs depict pointwise posterior means and 95% credibility regions of marginal posterior densities. The graphs in the left column depict results derived with the full sample of households and the years 2001 to 2005. The sample consists of 5,263 purchasing observations of 775 households; the results are equal to those depicted in figure I.3. The graphs in the right column depict results derived with a subsample of households and the years 2003 to 2005. The subsample consists of 1,441 purchasing observations of 242 households. All results were derived with the base prior specification (detailed in appendix A.1).

Appendix B

Information Disclosure in Open Non-Binding Procurement Auctions

B.1 Illustration: No information structure dominates the other

As the firms' first order conditions given in (II.3) and (II.6) are transcendental given any standard assumption about the distribution of the error terms ϵ_i, it is impossible to derive closed form solutions for the equilibrium prices in both the information and the no information case. In order to demonstrate that no information structure weakly dominates the other we thus resort to the use of numerical simulations.

We look at an auction with two bidding firms. The costs of the firms are $\mathbf{c} = (c_1, c_2) = (0, 1)$. We make the assumption that the error terms ϵ_i are iid type I extreme value distributed, and that the distribution of quality $f(q_j)$ is discrete: q_1 shall be drawn with probability 0.1, q_2 with probability 0.9.

Then for $\mathbf{q} = (q_1, q_2) = (0, 1)$ we get $\mathrm{EU} - \widetilde{\mathrm{EU}} = 0.75$. Thus, for these parameter values the buyer prefers the information case over the no information case. In contrast, for $\mathbf{q} = (0, 3)$ we get $\mathrm{EU} - \widetilde{\mathrm{EU}} = -0.34$. With these parameter values the buyer prefers the no information case over the information case.

B.2 Derivation of analytical results

We assume $\epsilon_2 - \epsilon_1$ to follow a uniform distribution with mean zeor and variance ν, and $\tilde{\epsilon}_2 - \tilde{\epsilon}_1$ to follow a uniform distribution with mean zero and variance $\tilde{\nu}$. It holds that $\tilde{\nu} \geq \nu$. Accordingly, the cumulative distribution function of $\epsilon_2 - \epsilon_1$ is given as

$$F_{\epsilon_2-\epsilon_1}(x) = \begin{cases} 0 & \text{for } x < -\sqrt{12\nu} \\ \frac{1}{2} + \frac{1}{\sqrt{12\nu}}x & \text{for } -\sqrt{12\nu} \leq x < \sqrt{12\nu} \\ 1 & \text{for } x \geq \sqrt{12\nu}, \end{cases} \tag{B.1}$$

and that of $\tilde{\epsilon}_2 - \tilde{\epsilon}_1$ as

$$F_{\tilde{\epsilon}_2-\tilde{\epsilon}_1}(x) = \begin{cases} 0 & \text{for } x < -\sqrt{12\tilde{\nu}} \\ \frac{1}{2} + \frac{1}{\sqrt{12\tilde{\nu}}}x & \text{for } -\sqrt{12\tilde{\nu}} \leq x < \sqrt{12\tilde{\nu}} \\ 1 & \text{for } x \geq \sqrt{12\tilde{\nu}}. \end{cases} \tag{B.2}$$

For the sake of exposition in the following we focus on the parameter space for which we get interior solutions. That is the parameter space for which both $-\sqrt{12\nu} \leq p_2^* - q_2 - p_1^* - q_1 < \sqrt{12\nu}$ and $-\sqrt{12\tilde{\nu}} \leq \tilde{p}_2^* - \tilde{p}_1^* < \sqrt{12\tilde{\nu}}$. p_i^* and $\tilde{p}_i^*$ are the equilibrium prices in the information respectively the no information case. These conditions hold if $0 \leq c_2 - c_1 < 3\sqrt{12\tilde{\nu}}$ and $-3\sqrt{12\nu} + (c_2 - c_1) < q_2 - q_1 \leq 3\sqrt{12\nu} + (c_2 - c_1)$. Note that for the parameter space depicted in figure II.1 we get interior solutions. In the complementary parameter space the situation in at least one of the two information cases turns deterministic, as due to the limited support of $\epsilon_2 - \epsilon_1$ respectively $\tilde{\epsilon}_2 - \tilde{\epsilon}_1$ randomness in the error terms no longer has an effect on the buyer's decision (as perceived by the bidders). This alters the position of the buyer's indifference line but has no effect on our basic finding that for large cost differences and small quality differences the buyer prefers the information case, while for small cost differences and large quality differences he prefers the no information case.

The firms' winning probabilities in the information case are

$$\begin{aligned} P_1(\mathbf{p}, \mathbf{q}) &= P(\epsilon_2 - \epsilon_1 \leq p_2 - q_2 - p_1 + q_1) = F_{\epsilon_2-\epsilon_1}(p_2 - q_2 - p_1 + q_1), \\ P_2(\mathbf{p}, \mathbf{q}) &= P(\epsilon_2 - \epsilon_1 > p_2 - q_2 - p_1 + q_1) = 1 - F_{\epsilon_2-\epsilon_1}(p_2 - q_2 - p_1 + q_1). \end{aligned}$$

If the P_j in the first order conditions (II.3) are expressed using the approximation (B.2), it is straightforward to solve these systems after the equilibrium prices $\mathbf{p}^*$:

$$\begin{aligned} p_1^* &= \frac{1}{3}(2c_1 + c_2) - \frac{1}{3}(q_2 - q_1) + \sqrt{3\nu}, \\ p_2^* &= \frac{1}{3}(c_1 + 2c_2) + \frac{1}{3}(q_2 - q_1) + \sqrt{3\nu}. \end{aligned}$$

The firms' winning probabilities in the no information case are

$$\begin{aligned}\tilde{P}_1(\mathbf{p},\mathbf{q}) &= P(\tilde{\epsilon}_2 - \tilde{\epsilon}_1 \leq \tilde{p}_2 - \tilde{p}_1) = F_{\tilde{\epsilon}_2-\tilde{\epsilon}_1}(\tilde{p}_2 - \tilde{p}_1),\\ \tilde{P}_2(\mathbf{p},\mathbf{q}) &= P(\tilde{\epsilon}_2 - \tilde{\epsilon}_1 > \tilde{p}_2 - \tilde{p}_1) = 1 - F_{\tilde{\epsilon}_2-\tilde{\epsilon}_1}(\tilde{p}_2 - \tilde{p}_1).\end{aligned}$$

Using the first order conditions (II.6), it follows that the equilibrium prices in the no information case are given as

$$\begin{aligned}\tilde{p}_1^* &= \frac{1}{3}(2c_1 + c_2) + \sqrt{3\tilde{\nu}},\\ \tilde{p}_2^* &= \frac{1}{3}(c_1 + 2c_2) + \sqrt{3\tilde{\nu}}.\end{aligned}$$

From simply comparing (p_1^*, p_2^*) to $(\tilde{p}_1^*, \tilde{p}_2^*)$, it follows that

$$\begin{aligned}p_1^* &= \tilde{p}_1^* - \frac{1}{3}(q_2 - q_1) - \sqrt{3}(\sqrt{\tilde{\nu}} - \sqrt{\nu}),\\ p_2^* &= \tilde{p}_2^* + \frac{1}{3}(q_2 - q_1) - \sqrt{3}(\sqrt{\tilde{\nu}} - \sqrt{\nu}).\end{aligned}$$

According to Small and Rosen (1981) the change in the buyer's expected utility from a change in the information structure can be computed as

$$\Delta\mathrm{EU} = \mathrm{EU} - \widetilde{\mathrm{EU}} = \int_{(\widetilde{W}_1,\widetilde{W}_2)}^{(W_1,W_2)} \left\{P_1(W_1,W_2)\mathrm{d}W_1 + [1 - P_1(W_1,W_2)]\mathrm{d}W_2\right\},$$

where $W_1 = q_1 - p_1$, $W_2 = q_2 - p_2$, $(\widetilde{W}_1, \widetilde{W}_2) = (q_1 - \tilde{p}_1^*, q_2 - \tilde{p}_2^*)$, $(W_1, W_2) = (q_1 - p_1^*, q_2 - p_2^*)$ and $P_1(W_1, W_2) = \frac{1}{2} + \sqrt{12\nu}(W_1 - W_2)$. Some algebra delivers

$$\begin{aligned}\mathrm{EU} - \widetilde{\mathrm{EU}} = {} & \frac{1}{3\sqrt{12\nu}}(q_2 - q_1)\left[(c_2 - c_1) - 2(q_2 - q_1)\right]\\ & +3(2\sqrt{\nu\tilde{\nu}} + \tilde{\nu} - 3\nu)\\ & +(\frac{\sqrt{\tilde{\nu}}}{2\sqrt{\nu}} - \frac{1}{2})(c_2 + c_1 - q_2 - q_1),\end{aligned}$$

as stated in the main body of the text.

Appendix C

Exploring the Opaqueness of the Patent System - Evidence from a Natural Experiment

C.1 Solution of the model

Table C.2 displays the normal form of the signaling game for information structures "public" and "private". We look for all Perfect Bayesian equilibria of these games which satisfy the "intuitive criterion" of Cho and Kreps (1987). The Perfect Bayesian equilibrium concept is a refinement of the Bayesian Nash equilibrium concept in the context of dynamic games with incomplete information, and Bayesian Nash equilibria can be deduced from the normal form of dynamic games. In practice, the determination of Bayesian Nash equilibria from the normal form of dynamic games is based on payoff comparisons.

In the following we first establish relationships between the payoffs of firm and between that of firm B. These relationships turn out to be dependent on specific parameters. As a second step we therefore divide the π_h^a-p-θ parameter space into subsets where the payoff relationships are non-ambiguous. Third, for each of these subsets we then derive all Bayesian Nash equilibria. Fourth, we check for every Bayesian Nash equilibrium whether if fulfills the criteria of a Perfect Bayesian Nash equilbrium (Bayesian beliefs and sequential rationality). Fifth, We check for every Perfect Bayesian Nash equilibrium whether it satisfies the "intuitive criterion". For reasons of brevity, we will describe steps three to five exemplarily for one subset of the π_h^a-p-θ parameter space only. The approach for all other subsets is completely analogous.

■ **Relationships between payoffs.** The first step in solving our game for both information structures is to find all Bayesian Nash equilibria. Essentially, the search for Bayesian Nash equilibria can be reduced to simple payoff comparisons, and these payoff comparisons can be traced back to comparisons of the payoffs in case both the value of firm A's patent and firm A's acceleration decision are public knowledge. We denote this information structure by "full". Note that information structure "full" describes the situation where the patent system is transparent with respect to patent value (and firm A's acceleration decision). The normal forms for information structure "full" are given in table C.1. We start with comparing the payoffs of the signaling game for information structure "full".

$\theta = h$	(o, o)		$(o, \neg o)$		$(\neg o, o)$		$(\neg o, \neg o)$	
a	$p\pi_h^a - c_a - c_o$	$-p\pi_h^a - c_o$	$p\pi_h^a - c_a - c_o$	$-p\pi_h^a - c_o$	$\pi_h^a - c_a$	$-\pi_h^a$	$\pi_h^a - c_a$	$-\pi_h^a$
$\neg a$	$p\pi_h^{\neg a} - c_o$	$-p\pi_h^{\neg a} - c_o$	$\pi_h^{\neg a}$	$-\pi_h^{\neg a}$	$p\pi_h^{\neg a} - c_o$	$-p\pi_h^{\neg a} - c_o$	$\pi_h^{\neg a}$	$-\pi_h^{\neg a}$

$\theta = l$	(o, o)		$(o, \neg o)$		$(\neg o, o)$		$(\neg o, \neg o)$	
a	$p\pi_l - c_a - c_o$	$-p\pi_l - c_o$	$p\pi_l - c_a - c_o$	$-p\pi_l - c_o$	$\pi_l - c_a$	$-\pi_l$	$\pi_l - c_a$	$-\pi_l$
$\neg a$	$p\pi_l - c_o$	$-p\pi_l - c_o$	π_l	$-\pi_l$	$p\pi_l - c_o$	$-p\pi_l - c_o$	π_l	$-\pi_l$

Table C.1: **Payoffs for full information.** The upper table shows the payoffs for full information in case $\theta = h$. The lower table shows the payoffs for full information in case $\theta = l$.

For firm A if the patent is of high value ($\theta = h$) the payoff comparisons are

$$p\pi_h^a - c_a - c_o \quad \text{vs.} \quad p\pi_h^{\neg a} - c_o, \tag{C.1}$$

$$p\pi_h^a - c_a - c_o \quad \text{vs.} \quad \pi_h^{\neg a}, \tag{C.2}$$

$$\pi_h^a - c_a \quad \text{vs.} \quad p\pi_h^{\neg a} - c_o, \tag{C.3}$$

$$\pi_h^a - c_a \quad \text{vs.} \quad \pi_h^{\neg a}, \tag{C.4}$$

and if the patent is of low value ($\theta = l$) the comparisons are

$$p\pi_l - c_a - c_o \quad \text{vs.} \quad p\pi_l - c_o, \tag{C.5}$$

$$p\pi_l - c_a - c_o \quad \text{vs.} \quad \pi_l, \tag{C.6}$$

$$\pi_l - c_a \quad \text{vs.} \quad p\pi_l - c_o, \tag{C.7}$$

$$\pi_l - c_a \quad \text{vs.} \quad \pi_l. \tag{C.8}$$

For firm B the payoff comparisons if the patent is of high value ($\theta = h$) are

$$-p\pi_h^a - c_o \quad \text{vs.} \quad -\pi_h^a, \tag{C.9}$$

$$-p\pi_h^{\neg a} - c_o \quad \text{vs.} \quad -\pi_h^{\neg a}, \tag{C.10}$$

and if the patent is of low value ($\theta = l$) the comparions are

$$-p\pi_l - c_o \quad \text{vs.} \quad -\pi_l. \tag{C.11}$$

The relationship between the payoffs in C.3 to C.8 is directly determined by our assumptions A1 to A4. For each other comparison there exists a certain cut-off value $p^{(\cdot)}$ at which the payoffs are equal. For all p smaller respectively larger then these $p^{(\cdot)}$ there exists a clear relationship between the underlying payoffs which follows directly from our assumptions A1 to A4. With "$\cdot|\cdot$" denoting the relationship left respectively right of the cut-off value $p^{(\cdot)}$, we have for firm A if the patent is of high value ($\theta = h$)

$$\begin{aligned}
p\pi_h^a - c_a - c_o &\lessgtr p\pi_h^{\neg a} - c_o, \quad \text{defines } p_1^A,\ <|>,\\
p\pi_h^a - c_a - c_o &\lessgtr \pi_h^{\neg a}, \quad \text{defines } p_2^A,\ <|>,\\
\pi_h^a - c_a &> p\pi_h^{\neg a} - c_o,\\
\pi_h^a - c_a &> \pi_h^{\neg a}.
\end{aligned}$$

For firm A if the patent is of low value ($\theta = l$) we have

$$\begin{aligned}
p\pi_l - c_a - c_o &< p\pi_l - c_o,\\
p\pi_l - c_a - c_o &< \pi_l,\\
\pi_l - c_a &> p\pi_l - c_o,\\
\pi_l - c_a &< \pi_l.
\end{aligned}$$

For firm B if the patent is of high value ($\theta = h$) we have

$$\begin{aligned}
-p\pi_h^a - c_o &\lessgtr -\pi_h^a, \quad \text{defines } p_3^B,\ >|<,\\
-p\pi_h^{\neg a} - c_o &\lessgtr -\pi_h^{\neg a} \quad \text{defines } p_2^B,\ >|<,
\end{aligned}$$

and if the patent is of low value ($\theta = l$)

$$-p\pi_l - c_o \lessgtr -\pi_l, \quad \text{defines } p_1^B,\ >|<.$$

Public	(o,o)		$(o,\neg o)$		$(\neg o,o)$		$(\neg o,\neg o)$	
(a,a)	$(1-\theta)p\pi_l + \theta p\pi_h^a - c_a - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^a - c_o$	$(1-\theta)p\pi_l + \theta p\pi_h^a - c_a - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^a - c_o$	$(1-\theta)\pi_l+\theta\pi_h^a- c_a$	$-(1-\theta)\pi_l-\theta\pi_h^a$	$(1-\theta)\pi_l+\theta\pi_h^a- c_a$	$-(1-\theta)\pi_l-\theta\pi_h^a$
$(a,\neg a)$	$(1-\theta)p\pi_l + \theta p\pi_h^a - \theta c_a - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^a - c_o$	$(1-\theta)\pi_l] + \theta p\pi_h^a - \theta c_a - \theta c_o$	$-(1-\theta)\pi_l - \theta p\pi_h^a - \theta c_o$	$(1-\theta)p\pi_l + \theta\pi_h^a - (1-\theta)c_o - \theta c_a$	$-(1-\theta)p\pi_l - \theta\pi_h^a - (1-\theta)c_o$	$(1-\theta)\pi_l]+\theta\pi_h^a- \theta c_a$	$-(1-\theta)\pi_l-\theta\pi_h^a$
$(\neg a,a)$	$(1-\theta)p\pi_l + \theta p\pi_h^{\neg a} - (1-\theta)c_a - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^{\neg a} - c_o$	$(1-\theta)p\pi_l + \theta\pi_h^{\neg a} - (1-\theta)c_a - (1-\theta)c_o$	$-(1-\theta)p\pi_l - \theta\pi_h^{\neg a} - (1-\theta)c_o$	$(1-\theta)\pi_l + \theta p\pi_h^{\neg a} - (1-\theta)c_a - \theta c_o$	$-(1-\theta)\pi_l - \theta p\pi_h^{\neg a} - \theta c_o$	$(1-\theta)\pi_l + \theta\pi_h^{\neg a} - (1-\theta)c_a$	$-(1-\theta)\pi_l - \theta\pi_h^{\neg a}$
$(\neg a,\neg a)$	$(1-\theta)p\pi_l + \theta p\pi_h^{\neg a} - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^{\neg a} - c_o$	$(1-\theta)\pi_l + \theta\pi_h^{\neg a}$	$-(1-\theta)\pi_l - \theta\pi_h^{\neg a}$	$(1-\theta)p\pi_l + \theta p\pi_h^{\neg a} - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^{\neg a} - c_o$	$(1-\theta)\pi_l + \theta\pi_h^{\neg a}$	$-(1-\theta)\pi_l] - \theta\pi_h^{\neg a}$

Private	o		$\neg o$	
(a,a)	$(1-\theta)p\pi_l + \theta p\pi_h^a - c_a - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^a - c_o$	$(1-\theta)\pi_l+\theta\pi_h^a- c_a$	$-(1-\theta)\pi_l-\theta\pi_h^a$
$(a,\neg a)$	$(1-\theta)p\pi_l + \theta p\pi_h^a - \theta c_a - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^a - c_o$	$(1-\theta)\pi_l]+\theta\pi_h^a- \theta c_a$	$-(1-\theta)\pi_l-\theta\pi_h^a$
$(\neg a,a)$	$(1-\theta)p\pi_l + \theta p\pi_h^{\neg a} - (1-\theta)c_a - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^{\neg a} - c_o$	$(1-\theta)\pi_l + \theta\pi_h^{\neg a} - (1-\theta)c_a$	$-(1-\theta)\pi_l - \theta\pi_h^{\neg a}$
$(\neg a,\neg a)$	$(1-\theta)p\pi_l + \theta p\pi_h^{\neg a} - c_o$	$-(1-\theta)p\pi_l - \theta p\pi_h^{\neg a} - c_o$	$(1-\theta)\pi_l + \theta\pi_h^{\neg a}$	$-(1-\theta)\pi_l - \theta\pi_h^{\neg a}$

Table C.2: **Normal form of the game for information structures "public" and "private".** For both information structures each row represents a possible strategy of firm A, while each column represents a possible strategy of firm B. For each strategy of firm A its actions are conditional on nature's draw of the patent value. That is, (a,a) is short for $(a|v=h, a|v=l)$, and so on. For information structure "public", for each strategy of firm B its actions are conditional on whether firm B observes accelerated patent examination. That is, (o,o) is short for $(o|a, o|\neg a)$. For each information structure each box displays the payoffs of firm A (left) and firm B (right) if the respective strategies are played.

The cut-off values are defined as follows:

$$\begin{aligned}
p_1^A &= \frac{c_a}{\pi_h^a - \pi_h^{\neg a}}, \\
p_2^A &= \frac{c_a + c_o + \pi_h^{\neg a}}{\pi_h^a}, \\
p_1^B &= \frac{\pi_l - c_o}{\pi_l}, \\
p_2^B &= \frac{\pi_h^{\neg a} - c_o}{\pi_h^{\neg a}}, \\
p_3^B &= \frac{\pi_h^a - c_o}{\pi_h^a}.
\end{aligned}$$

The payoffs for information structures “public” and “private” are composed from the payoffs for information structure “full”. Thus, with information about the relationships between the payoffs for information structure “full” it is easy to derive the relationships between the payoffs for information structures “public” and “private”. Each row in table C.2 corresponds to a strategy of firm A, and each column to a strategy of firm B. First, we determine the best reactions of firm A to each possible strategy of firm B. Based on our results for the payoffs of firm A for information structure “full” and our assumptions A1 to A4 we find for information structure “public”:

1st column: 4th row if $p < p_1^A$, 2nd row if $p > p_1^A$.

2nd column: 4th row if $p < p_2^A$, 2nd row if $p > p_2^A$.

3rd column: 1st row.

4th column: 2nd row.

The results for firm A and Information structure “private” are:

1st column: 4th row if $p < p_1^A$, 2nd row if $p > p_1^A$.

2nd column: 2nd row.

With that, the relationships between the payoffs of firm A are fully determined.

The results for firm B and information structure “public” are:

2nd row: 1st column if $0 < p < p_1^B$, 2nd column if $p_1^B < p < p_3^B$, 4th column if $p_3^B < p < 1$.

3rd row: 1st column if $0 < p < p_1^B$, 3rd column if $p_1^B < p < p_2^B$, 4th column if $p_2^B < p < 1$.

For the 1st row of information structure "public" and the 1st and 2nd row of information structure "private" the same payoffs have to be compared. The comparison to be made is

$$-(1-\theta)p\pi_l - \theta p\pi_h^a - c_o \text{ vs. } -(1-\theta)\pi_l - \theta\pi_h^a.$$

The relationship between these payoffs depends on the relationship between p and θ. With

$$p_{\theta,1} = 1 - \frac{c_o}{\theta\pi_h^a + (1-\theta)\pi_l}, \tag{C.12}$$

we have equality for $p = p_{\theta,1}$. For values of p smaller than $p_{\theta,1}$ the former payoff is larger than the latter, and vice versa. For $\theta = 0$ $p_{\theta,1}$ equals p_1^B, and for $\theta = 1$ $p_{\theta,1}$ equals p_3^B. We denote the inverse function of $p_{\theta,1}(\theta)$ by $\theta_1(p)$. The situation for the 4th row of information structure "public" and the 3rd and 4th row of information structure "private" is analogous: The comparison to be made is

$$-(1-\theta)p\pi_l - \theta p\pi_h^{\neg a} - c_o \text{ vs. } -(1-\theta)\pi_l - \theta\pi_h^{\neg a}.$$

The relationship between these payoffs depends on the relationship between p and θ. With

$$p_{\theta,2} = 1 - \frac{c_o}{\theta\pi_h^{\neg a} + (1-\theta)\pi_l}, \tag{C.13}$$

we have equality for $p = p_{\theta,2}$. For values of p smaller than $p_{\theta,2}$ the former payoff is larger than the latter, and vice versa. For $\theta = 0$ $p_{\theta,2}$ equals p_1^B, and for $\theta = 1$ $p_{\theta,2}$ equals p_2^B. We denote the inverse function of $p_{\theta,2}(\theta)$ by $\theta_2(p)$. With that we can complete the payoff comparisons for firm B. For information structure "public" we have:

1st row: 1st and 2nd column if $< p_{\theta,1}$, 3rd and 4th column if $p > p_{\theta,1}$.

4th row: 1st and 3rd column if $p < p_{\theta,2}$, 2nd and 4th column if $p > p_{\theta,2}$.

For information structure "private" the results are:

1st row: 1st column if $< p_{\theta,1}$, 2nd column if $p > p_{\theta,1}$.

2nd row: 1st column if $< p_{\theta,1}$, 2nd column if $p > p_{\theta,1}$.

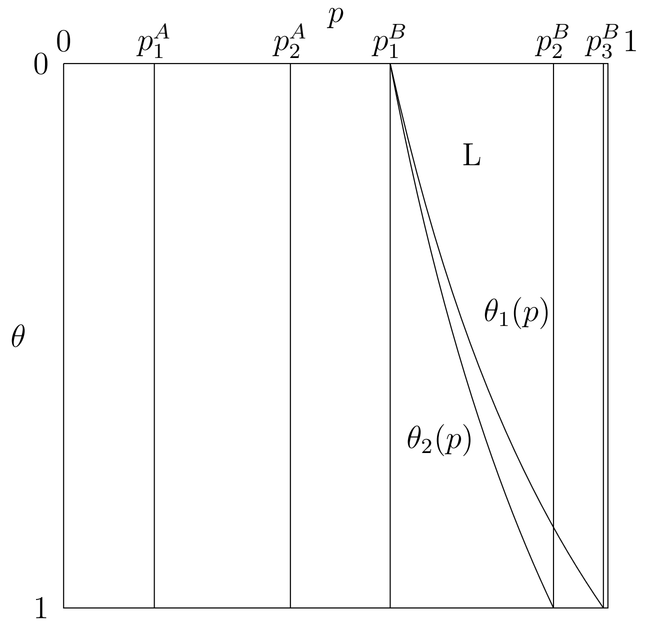

Figure C.1: **p-θ subsets for the π_h^a subset Π_6.**

3rd row: 1st column if $< p_{\theta,2}$, 2nd column if $p > p_{\theta,2}$.

4th row: 1st column if $p < p_{\theta,2}$, 2nd column if $p > p_{\theta,2}$.

With that, the relationships between the payoffs of firm B are fully determined.

■ **Subsets of the π_h^a-p-θ parameter space.** From assumptions A1 to A4 it follows that $p_1^A < p_2^A$ and $p_1^B < p_2^B < p_3^B$. The relationship between the boundaries of firm A (p_1^A, p_2^A) and that of firm B (p_1^B, p_2^B, p_3^B) depends on the value of π_h^a. We can define different subsets $\Pi_{(\cdot)}$ for π_h^a:

$$\begin{aligned}
\Pi_1: &\quad c_a + \pi_h^{\neg a} < \pi_h^a < \frac{\pi_l}{\pi_l - c_o} c_a + \pi_h^{\neg a} \\
\Pi_2: &\quad \frac{\pi_l}{\pi_l - c_o} c_a + \pi_h^{\neg a} < \pi_h^a < c_a + c_o + \pi_h^{\neg a} \\
\Pi_3: &\quad c_a + c_o + \pi_h^{\neg a} < \pi_h^a < c_a + 2c_o + \pi_h^{\neg a} \\
\Pi_4: &\quad c_a + 2c_o + \pi_h^{\neg a} < \pi_h^a < \frac{\pi_h^{\neg a}}{\pi_h^{\neg a} - c_o} [c_a + c_o + \pi_h^{\neg a}] \\
\Pi_5: &\quad \frac{\pi_h^{\neg a}}{\pi_h^{\neg a} - c_o} [c_a + c_o + \pi_h^{\neg a}] < \pi_h^a < \frac{\pi_l}{\pi_l - c_o} [c_a + c_o + \pi_h^{\neg a}] \\
\Pi_6: &\quad \frac{\pi_l}{\pi_l - c_o} [c_a + c_o + \pi_h^{\neg a}] < \pi_h^a
\end{aligned}$$

For each subset $\Pi_{(\cdot)}$ there follows a clear relationship between the boundaries of firm A (p_1^A, p_2^A) and that of firm B (p_1^B, p_2^B, p_3^B) from our assumptions A1 to A4 :

$$\begin{aligned}
\Pi_1: &\quad 0 < p_1^B < p_1^A < p_2^B < p_3^B < 1 \\
\Pi_2: &\quad 0 < p_1^A < p_1^B < p_2^B < p_3^B < 1 \\
\Pi_3: &\quad 0 < p_1^A < p_1^B < p_2^B < p_3^B < p_2^A < 1 \\
\Pi_4: &\quad 0 < p_1^A < p_1^B < p_2^B < p_2^A < p_3^B < 1 \\
\Pi_5: &\quad 0 < p_1^A < p_1^B < p_2^A < p_2^B < p_3^B < 1 \\
\Pi_6: &\quad 0 < p_1^A < p_2^A < p_1^B < p_2^B < p_3^B < 1
\end{aligned}$$

For each subset $\Pi_{(\cdot)}$, the curve $p_{\theta,1}$ runs from $(p = p_1^B, \theta = 0)$ to $(p = p_3^B, \theta = 1)$, and the curve $p_{\theta,2}$ from $(p = p_1^B, \theta = 0)$ to $(p = p_2^B, \theta = 1)$.

To this point we have separated the 3-dimensional π_h^a-p-θ parameter space into several subsets. Figure C.1 exemplarily displays the p-θ subsets for the π_h^a subset Π_6.

■ **Bayesian Nash equilibria.** A Bayesian Nash equilibrium is a pair of strategies for which firm A's strategy is a best response to firm B's strategy given his own type and his beliefs about firm B's type, and vice versa. A Bayesian Nash equilibrium can be interpreted as a Nash equilibrium of an expanded game, where the firms' pure strategies are type-contingent. Thus, a Bayesian Nash equilibrium is a pair of strategies of the expanded game for which firm A's strategy is a best response to firm B's strategy and vice versa.

The payoff matrices in table C.2 are payoff matrices of expanded games. Each possible strategy of firm A is represented by a row, and each possible strategy of firm B by a column. In order to determine Bayesian Nash equilibria, we have to determine the best reaction of firm A to each strategy of firm B and vice versa. In practice, that means for each column of the matrices in table C.2 we first have to find the row with the highest payoff for firm A (respectively for each row the column with the highest payoff for firm B). A Bayesian Nash equilibrium then corresponds to a cell in the output matrix for information structure "public" (respectively to a cell in the output matrix for information structure "private") which contains both the highest payoff of firm A in the respective column and the highest payoff of firm B in the respective row.

As the relationships between the payoffs depend on which subset of the π_h^a-p-θ parameter space we are in, we have to determine Bayesian Nash equilibria separately for every subset of the π_h^a-p-θ space. The procedure thereby is always the same. Thus, for reasons of brevity we will exemplarily demonstrate the determination of Bayesian Nash equilibria (and the subsequent determination of Perfect Bayesian Nash equilibria and the application of the intuitive criterion) for one subset of our parameter space. We marked this subset as subset "L" in figure C.1.

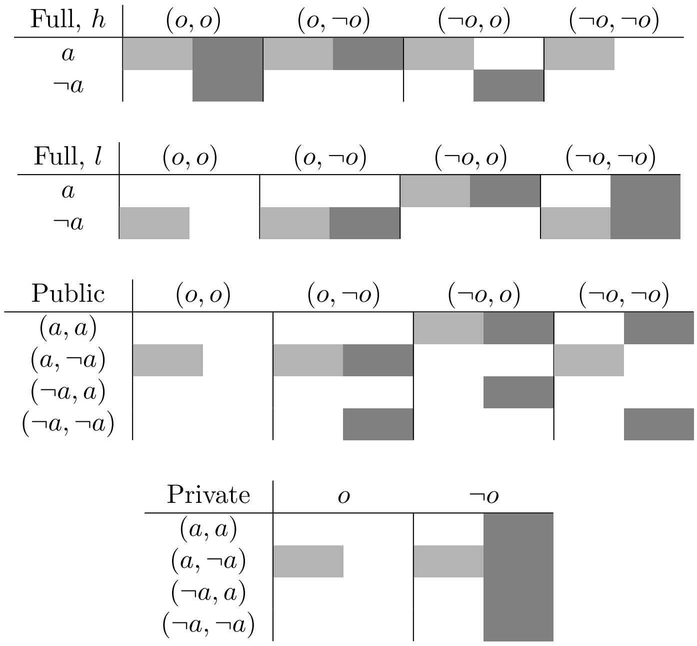

Table C.3: **Bayesian Nash equilibria.** Displayed are schematic payoff matrices for information structures "full", "public" and "private" and subset "L" of the parameter space (see figure C.1). In each matrix the highest payoffs of firm A in each column and of firm B in each row are marked. Bayesian Nash equilibria are cells which contain both the highest payoff of firm A and firm B.

In the schematic payoff matrices in table C.3 the highest payoffs of firm A in each column and of firm B in each row are marked for each payoff structure. Bayesian Nash equilibria are cells which contain both the highest payoff of firm A and firm B. In case the patent system is transparent and the patent is of high value there are two Bayesian Nash equilibria: $[a;(o,o)]$ and $[a;(o,\neg o)]$. In case the patent system is transparent and the patent is of low value there are three equilibria: $[\neg a;(o,\neg o)]$, $[a;(\neg o,o)]$ and $[\neg a;(\neg o,\neg o)]$. For information structure "public" there are two equilibria: $[(a,\neg a);(o,\neg o)]$ and $[(a,a);(\neg o,o)]$. For information structure "private" there is one equilibrium: $[(a,\neg a);\neg o]$.

■ **Perfect Bayesian Nash equilibria.** For information structures "public" and "private" we check for every Bayesian Nash equilibrium whether it fulfills the criteria of a Perfect Bayesian Nash equilibrium - that is, whether there is a belief structure which is consistent with this equilibrium. We exemplarily demonstrate the procedure for the two equilibria of information structure "public" in subset "L".

For the separating equilibrium $[(a, \neg a); (o, \neg o)]$ a belief structure of firm B which is consistent with this equilibrium is as follows: Firm B puts probability one on the event "firm A's patent is of high value" if it observes acceleration. If it does not observe acceleration it puts probability one on the event "firm A's patent is of low value". It is easy to show that this belief is consistent with the equilibrium: If firm B believes that firm A's patent is of high value, it is optimal for firm B to oppose firm A's patent. The reason is that for $p < p_3^B = \frac{\pi_h^a - c_o}{\pi_h^a}$ (which is the case for subset "L") firm B's payoff in case it opposes an accelerated high-value patent of firm A $(-p\pi_h^a - c_o)$ is larger than its payoff in case it does not oppose $(-\pi_h^a)$. If firm B believes that firm A's patent is of low value it is optimal for firm B not to oppose firm A's patent. The reason is that for $p > p_1^B = \frac{\pi_l - c_o}{\pi_l}$ (which is the case for subset "L") firm B's payoff in case it does not oppose a non-accelerated low-value patent of firm A $(-\pi_l)$ is larger than its payoff in case it does not oppose $(-p\pi_l - c_o)$. Given that firm B opposes an accelerated patent, firm A only benefits from accelerating a high-value patent. The reason is that in subset "L" in case firm A has a low-value patent its payoff in case it does not accelerate the patent (π_l) is obviously larger than its payoff in case it accelerates the patent $(p\pi_l - c_o)$. (Note that $p < 1$.) In contrast, in case firm A has a high-value patent its payoff in case it accelerates the patent $(p\pi_h^a - c_o - c_a)$ is larger than its payoff in case it does not accelerate the patent $(\pi_h^{\neg a})$. (In subset "L" it holds that $p > p_2^A = \frac{\pi_h^{\neg a} + c_a + c_o}{\pi_h^a}$. Thus, in subset "L" it holds that $p\pi_h^a - c_o - c_a \geq \pi_h^{\neg a}$.)

For the pooling equilibrium $[(a, a); (\neg o, o)]$ a belief structure of firm B which is consistent with this equilibrium is as follows: Off the equilibrium path firm B puts probability one on the event "firm A has a high-value patent". On the equilibrium path firm B puts probability θ on the event "firm A has a high-value patent" and probability $1 - \theta$ on the event "firm A has a low-value patent". Given this belief structure, on the equilibrium path firm B's payoff in case it does not oppose $((1 - \theta)(-\pi_l) + \theta(-\pi_h^a))$ is larger than its payoff in case it does oppose $((1-\theta)(-p\pi_l - c_o) + \theta(-p\pi_h^a - c_o))$. The reason is that in subset "L" it holds that $p > p_{\theta,1} = 1 - \frac{c_o}{\theta\pi_h^a + (1-\theta)\pi_l}$:

$$\begin{aligned} p &> 1 - \frac{c_o}{\theta\pi_h^a + (1-\theta)\pi_l} \quad \Leftrightarrow \\ c_o &> (1-p)\theta\pi_h^a + (1-p)(1-\theta)\pi_l \quad \Leftrightarrow \\ (1-\theta)(-\pi_l) + \theta(-\pi_h^a) &> (1-\theta)(-p\pi_l - c_o) + \theta(-p\pi_h^a - c_o) \end{aligned}$$

As in subset "L" it holds that $p < p_2^B = \frac{\pi_h^{\neg a} - c_o}{\pi_h^{\neg a}}$, off the equilibrium path firm B's payoff in case it opposes $(-p\pi_h^{\neg a})$ is larger than its payoff in case it does not oppose $(-\pi_h^{\neg a})$. Given that in case firm B would observe acceleration it would oppose, firm A is better off accelerating both high-value and low-value patents. The reason is that both for high-value patents and for low-value patents the payoff of firm A in case it does accelerate and firm B does not oppose is larger than its payoff in case it does not accelerate and firm B does oppose ($\pi_l - c_a > p\pi_l - c_o$ and $\pi_h^a - c_a > p\pi_h^{\neg a} - c_o$).

■ **Intuitive criterion.** For some subsets of the π_h^a-θ-p parameter space we find several Perfect Bayesian Nash equilibria for information structures "public" and "private". We use the "intuitive criterion" introduced by Cho and Kreps (1987) to reduce the number of equilibria. The "intuitive criterion" uses a forward induction argument: It eliminates equilibria when firm A would be better off if it deviated from the equilibrium. We demonstrate the use of the "intuitive criterion" exemplarily for information structure "public" and subset "L" of our parameter space. There we have two equilibria which fulfill the criteria of a Perfect Bayesian Nash equilibrium. These are the separating equilibrium $[(a, \neg a); (o, \neg o)]$ and the pooling equilibrium $[(a, a); (\neg o, o)]$. For the separating equilibrium there is no deviation which would make firm A better off. However, the pooling equilibrium fails the intuitive criterion:

For the pooling equilibrium $[(a, a); (\neg o, o)]$ to be sequentially rational firm B has to believe that firm A has a high-value patent if it does not accelerate. However, this belief is not plausible: If firm A has a high-value patent, in equilibrium it gets $\pi_h^a - c_a$. When firm A deviates, it only gets $\pi_h^{\neg a}$. Yet, if firm A has a low-value patent, it has an incentive to deviate: In equilibrium, firm A gets $\pi_l - c_a$ if it has a low-value patent. However, if firm A deviates and convinces firm B that it has a low-value patent, it gets π_l (because if convinced firm B would not oppose). Thus, firm B should put zero probability on firm A having a high-value patent when firm A does not accelerate. However, in this case firm B would play $\neg o$ in reaction to $\neg a$, which upsets the equilibrium. That is, the pooling equilibrium $[(a, a); (\neg o, o)]$ fails the intuitive criterion.

■ **Results.** We summarize our results in figure C.2 and table C.4. Figure C.2 displays all subsets of the π_h^a-θ-p parameter space with specific relationships between the payoffs of firm A and firm B. We marked these subsets by romanic upper-case letters. For each of these subsets and each information structure table C.4 displays all Perfect Bayesian Nash equilibria which fulfill the intuitive criterion introduced by Cho and Kreps (1987). Note that we did not display the subsets for very low

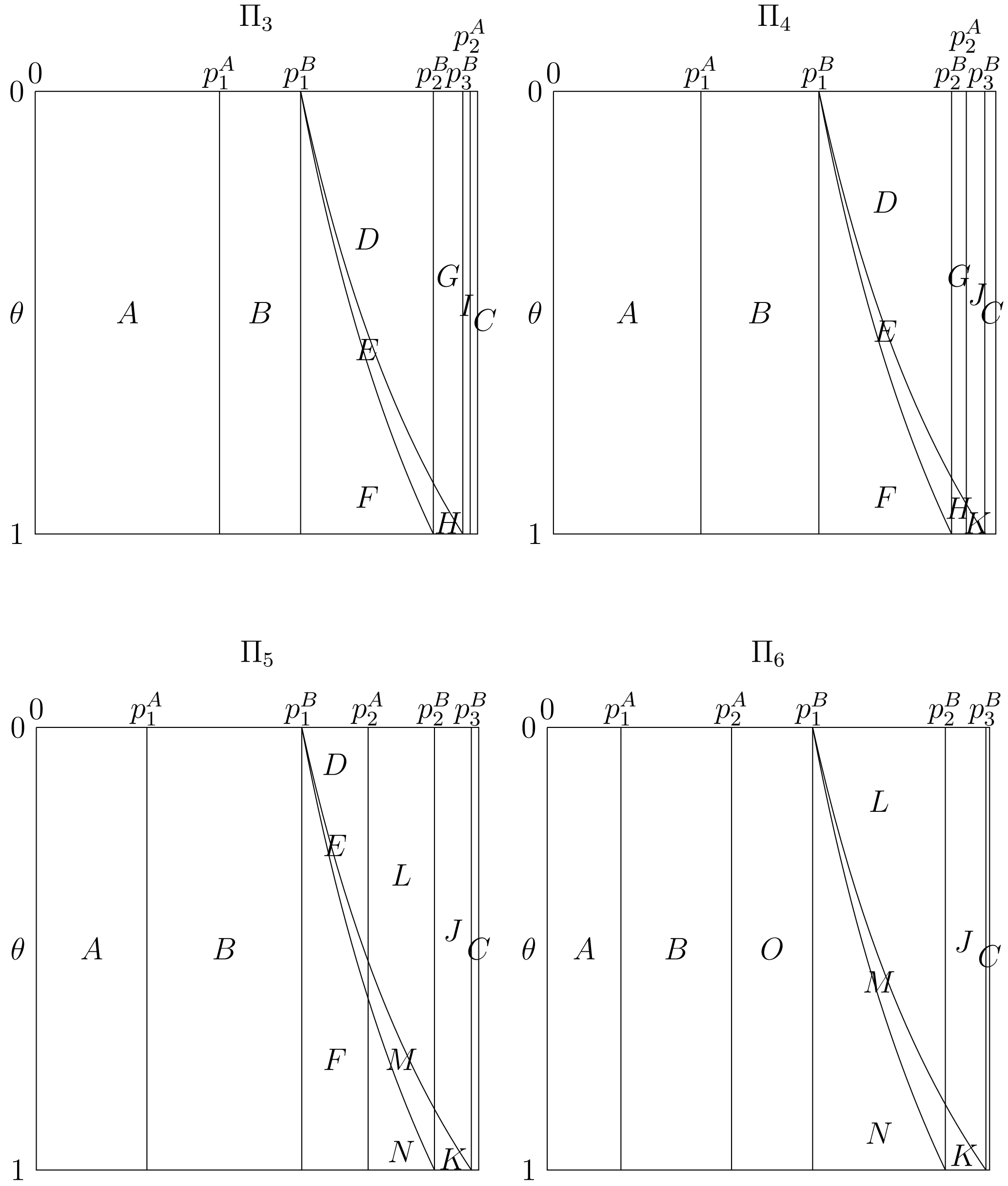

Figure C.2: **Subsets of the π_h^a-θ-p parameter space.** Each graph displays the complete θ-p parameter space for a subset $\Pi_{(\cdot)}$ of the π_h^a parameter space. We marked subsets of the π_h^a-θ-p parameter space with specific payoff relationships by romanic upper-case letters. For each of these subsets and each information structure all Perfect Bayesian Nash equilibria which fulfill the intuitive criterion are given in table C.4.

gains from acceleration (Π_1 and Π_2). The reason is that cases where there are no economically significant gains from acceleration are uninteresting for our analysis. From the equilibrium strategies of the firms it is easy to derive expected outcomes for each subset of the parameter space and each information structure. It shows that for some of the subsets marked in figure C.2 outcomes are the same for all

	A	**B**	**C**	**D**
Full, h	$\neg a; (o, o)$	$a; (o, o)$	$a; (\neg o, \neg o)$	$a; (o, o)$
Full, l	$\neg a; (o, o)$	$\neg a; (o, o)$	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$
Public	$(\neg a, \neg a); (o, o)$	$(a, \neg a); (o, o)$	$(a, \neg a); (\neg o, \neg o)$	$(\neg a, \neg a); (o, \neg o)$
Private	$(\neg a, \neg a); o$	$(a, \neg a); o$	$(a, \neg a); \neg o$	$(a, \neg a); \neg o$
	E	**F**	**G**	**H**
Full, h	$a; (o, o)$	$a; (o, o)$	$\neg a; (o, \neg o)$	$\neg a; (o, \neg o)$
Full, l	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$
Public	$(\neg a, \neg a); (o, \neg o)$	No eq.	$(\neg a, \neg a); (o, \neg o)$	$(\neg a, \neg a); (o, \neg o)$
Private	$(a, \neg a); o$	$(a, \neg a); o$	$(a, \neg a); \neg o$	$(a, \neg a); o$
	I	**J**	**K**	**L**
Full, h	$a; (\neg o, \neg o)$	$a; (o, \neg o)$	$a; (o, \neg o)$	$a; (o, o)$
Full, l	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$
Public	$(a, \neg a); (\neg o, \neg o)$	$(a, \neg a); (o, \neg o)$	$(a, \neg a); (o, \neg o)$	$(a, \neg a); (o, \neg o)$
Private	$(a, \neg a); \neg o$	$(a, \neg a); \neg o$	$(a, \neg a); o$	$(a, \neg a); \neg o$
	M	**N**	**O**	
Full, h	$a; (o, o)$	$a; (o, o)$	$a; (o, o)$	
Full, l	$\neg a; (\neg o, \neg o)$	$\neg a; (\neg o, \neg o)$	$\neg a; (o, o)$	
Public	$(a, \neg a); (o, \neg o)$	$(a, \neg a); (o, \neg o)$	$(a, \neg a); (o, o)$	
Private	$(a, \neg a); o$	$(a, \neg a); o$	$(a, \neg a); o$	

Table C.4: **Perfect Bayesian Nash equilibria which fulfill the intuitive criterion for all subsets of the π_h^a-θ-p parameter space and all information structures.** For each subset of the parameter space and each information structure firm A's strategy is given before the semicolon, and firm B's strategy is given after the semicolon. In case firm A's actions ("accelerate" or "not accelerate") are contingent on the draw of the patent value, the first entry in the parentheses gives firm A's action in case the patent is of high value, and the second entry gives firm A's action in case the patent is of low value. In case firm B's actions ("oppose" or "not oppose") are contingent on firm A's acceleration decision, the first entry in parentheses gives firm B's action in case firm A accelerates, and the second entry gives firm B's action in case firm A does not accelerate.

information structure. Thus, with respect to outcomes we can combine some of the subsets. In result we get the graphs in figures III.3 and III.4, which display the outcomes for every subset of the parameter space.

C.2 Welfare calculations

Table C.5 displays the differences in firms' expected payoffs between information structures "public", "private" and "full" for subsets I to IV (see figures III.3 and III.4). Information structure "full" denotes the case of full transparency of the patent system with respect to patent value (and firm A's acceleration decision). The differences are computed for each firm individually and for both firms in aggregation. In case firms are considered in aggregation table C.5 shows that whether the displayed differences in payoffs are positive or negative is clear from our assumptions A1 to A4. In case firms are considered individually also most of the payoff differences are of a clear sign. However, there are some payoff differences which critically depend on p and which we consider separately in the following:

	Firm A		Firm B	
Full I, II - Public I, II	$\theta[p\pi_h^a - \pi_h^{\neg a} - c_a - c_o]$	$(-)$	$\theta[\pi_h^{\neg a} - p\pi_h^a - c_o]$	$(+\vert-)$
Full III, IV - Public III, IV	0	(0)	0	(0)
Full I, III - Private I, III	$-\theta[(1-p)\pi_h^a + c_o]$	$(-)$	$\theta[(1-p)\pi_h^a - c_o]$	$(+)$
Full II, IV - Private II, IV	$(1-\theta)[(1-p)\pi_l + c_o]$	$(+)$	$(1-\theta)[-(1-p)\pi_l + c_o]$	$(+)$
Public I - Private I	$-\theta[\pi_h^a - \pi_h^{\neg a} - c_a]$	$(-)$	$\theta[\pi_h^a - \pi_h^{\neg a}]$	$(+)$
Public II - Private II	$\theta[\pi_h^{\neg a} - p\pi_h^a + c_a]$		$\theta[-\pi_h^{\neg a} + p\pi_h^a]$	
	$+c_o + (1-\theta)(1-p)\pi_l$	$(+)$	$+c_o - (1-\theta)(1-p)\pi_l$	$(-\vert+)$
Public III - Private III	$-\theta[(1-p)\pi_h^a + c_o]$	$(-)$	$\theta[(1-p)\pi_h^a - c_o]$	$(+)$
Public IV - Private IV	$(1-\theta)[(1-p)\pi_l + c_o]$	$(+)$	$(1-\theta)[-(1-p)\pi_l + c_o]$	$(+)$

	Firm A + Firm B	
Full I, II - Public I, II	$-\theta(c_a + 2c_o)$	$(-)$
Full III, IV - Public III, IV	0	(0)
Full I, III - Private I, III	$-2\theta c_o$	$(-)$
Full II, IV - Private II, IV	$(1-\theta)2c_o$	$(+)$
Public I - Private I	θc_a	$(+)$
Public II - Private II	$\theta c_a + 2c_o$	$(+)$
Public III - Private III	$-2\theta c_o$	$(-)$
Public IV - Private IV	$(1-\theta)2c_o$	$(+)$

Table C.5: **Payoff Differences.** Displayed are the differences in payoffs between subsets I to IV and information structures "public", "private" and "full" for firms A and B individually and considered together. Information structure "full" means full transparency of the patent system. The subsets I to IV are marked in figures III.3 and III.4. The payoff difference for firm B between information structures "full" and "public" and subsets I and II is positive if $p < \frac{\pi_h^{\neg a} - c_o}{\pi_h^a}$ and negative otherwise. The payoff difference for firm B between information structures "public" and "private" and subset II is negative if $p < \frac{\theta\pi_h^{\neg a} + (1-\theta)\pi_l - c_o}{\theta\pi_h^a + (1-\theta)\pi_l}$ and positive otherwise.

□ **Full I, II - Public I, II.** In subsets I and II it holds that $p < p_2^A = \frac{\pi_h^{\neg a} + c_a + c_o}{\pi_h^a}$. With that it follows directly that the payoff difference for firm A is negative. The sign of the payoff difference for firm B remains ambiguous: It is (weakly) positive if $p \leq \frac{\pi_h^{\neg a} - c_o}{\pi_h^a}$ and negative otherwise. Depending on the exact relationship between c_o, π_l and $\pi_h^{\neg a}$ the cut-off $\frac{\pi_h^{\neg a} - c_o}{\pi_h^a}$ lies either between p_1^B and p_2^A or is smaller than p_1^B. (Our assumptions A1 to A4 allow both.)

□ **Full II, IV - Private II, IV and Public IV - Private IV, firm B.** In subsets II and IV it holds that $p > p_1^B = \frac{\pi_l - c_o}{\pi_l}$. It follows that $1 - p < \frac{c_o}{\pi_l}$. With that it is clear that the payoff difference is positive.

□ **Public II - Private II.** In subset II it holds that $p < p_2^A = \frac{\pi_h^{\neg a} + c_a + c_o}{\pi_h^a}$. Given that relation and assumption A4 ($c_a \leq c_o$), it follows directly that the payoff difference for firm A is positive. Whether the payoff difference for firm B is positive or negative is not clear: It is (weakly) negative if $p \leq \frac{\theta\pi_h^{\neg a}+(1-\theta)\pi_l - c_o}{\theta\pi_h^a+(1-\theta)\pi_l}$ and positive otherwise. Depending on the exact relationship between c_o, π_l and $\pi_h^{\neg a}$ the cut-off value $p_W = \frac{\theta\pi_h^{\neg a}+(1-\theta)\pi_l - c_o}{\theta\pi_h^a+(1-\theta)\pi_l}$ lies either between $p_{\theta,1}$ and $p_{\theta,2}$ or is smaller than $p_{\theta,2}$. (Our assumptions allow both.)

C.3 Breakdown of applications by residency and technical field

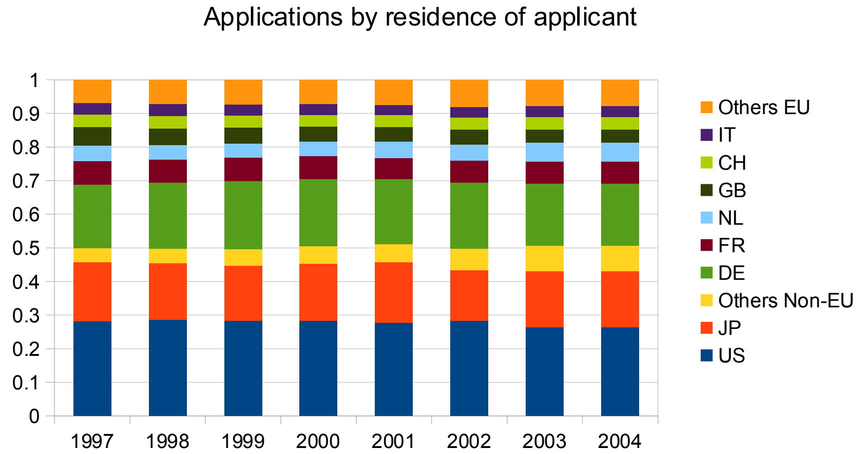

Figure C.3: **Breakdown of applications by residency.** Source: Own computations based on EPASYS data of the EPO.

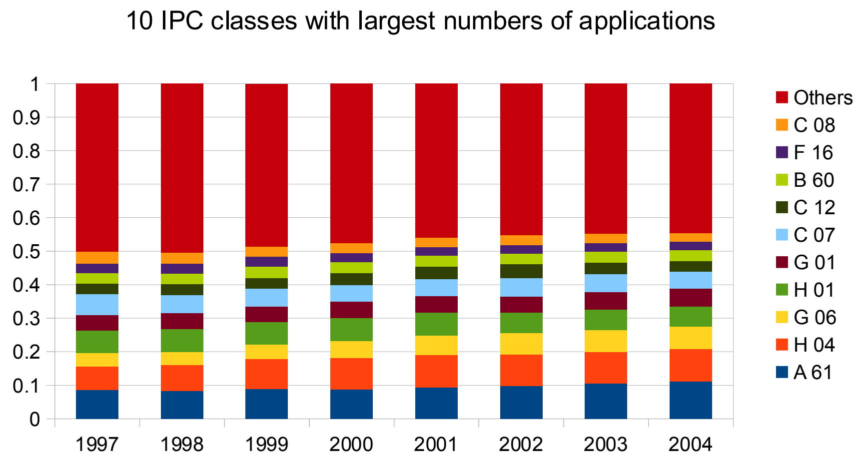

Figure C.4: **Breakdown of applications by IPC classes.** C08: Organic macromolecular compounds; F16: Engineering elements; B60: Vehicles in general; C12: Biochemistry; C07: Organic Chemistry; G01: Measuring, Testing; H01: Basic electric elements; G06: Computing; H04: Electric communication technique; A61: Medical or veterinary science. Source: Own computations based on EPASYS data of the EPO.

List of Figures

List of Tables

Bibliography

ANDERSSON, F. (2002): "Pooling reputations," *International Journal of Industrial Organization*, 20, 715–730.

ASKER, J. AND E. CANTILLON (2008): "Properties of scoring auctions," *RAND Journal of Economics*, 39, 69–85.

——— (2010): "Procurement when price and quality matter," *RAND Journal of Economics*, 41, 1–34.

ATHEY, S. AND J. LEVIN (2001): "Information and competition in U.S. Forest Service timber auctions," *Journal of Political Economy*, 109, 375–417.

BALACHANDER, S. AND S. GHOSE (2003): "Reciprocal Spillover Effects: A Strategic Benefit of Brand Extensions," *The Journal of Marketing*, 67, 4–13.

BESSEN, J. AND M. J. MEURER (2008): *Patent failure: How judges, bureaucrats, and lawyers put innovators at risk*, Princeton University Press, 1st ed.

BICHLER, M. AND J. KALAGNANAM (2005): "Configurable Offers and Winner Determination in Multi-Attribute Auctions," *European Journal of Operations Research*, 160, 380–394.

BRANCO, F. (1997): "The design of multidimensional auctions," *RAND Journal of Economics*, 28, 63–81.

BRONNENBERG, B. J., M. W. KRUGER, AND C. F. MELA (2008): "Database paper: The IRI marketing data set," *Marketing Science*, 27, 745–748.

CABRAL, L. M. B. (2000): "Stretching Firm and Brand Reputation," *The RAND Journal of Economics*, 31, 658–673.

——— (2009): "Umbrella branding with imperfect observability and moral hazard," *International Journal of Industrial Organization*, 27, 206–213.

Caplin, A. and B. Nalebuff (1991): "Aggregation and imperfect competition: on the existence of equilibrium," *Econometrica*, 59, 25–59.

Che, Y.-K. (1993): "Design competition through multidimensional auctions," *RAND Journal of Economics*, 24, 668–680.

Chen-Ritzo, C.-H., T. P. Harrison, A. M. Kwasnica, and D. J. Thomas (2005): "Better, faster, cheaper: An experimental analysis of a multiattribute reverse auction mechanism with restricted information feedback," *Management Science*, 51, 1753–1762.

Cho, I.-K. and D. M. Kreps (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102, 179–221.

Choi, J. P. (1998): "Brand Extension as Informational Leverage," *The Review of Economics Studies*, 65, 655–669.

Colucci, D., N. Doni, and V. Valori (2011): "Information disclosure in procurement auctions with horizontally differentiated suppliers," Working Paper.

Doni, N. and D. Menicucci (2010): "A note on informational disclosure in procurement auctions," *Economics Letters*, 108, 307–310.

Dube, J.-P., G. J. Hitsch, and P. E. Rossi (2010): "State Dependence and Alternative Explanations for Consumer Inertia," *The RAND Journal of Economics*, 41, 417–445.

Engelbrecht-Wiggans, R., E. Haruvy, and E. Katok (2007): "A comparison of buyer-determined and price-based multiattribute mechanisms," *Marketing Science*, 26, 629–641.

Erdem, T. (1998): "An Empirical Analysis of Umbrella Branding," *Journal of Marketing Research*, 35, 339–351.

Erdem, T., M. P. Keane, and B. Sun (2008): "The impact of advertising on consumer price sensitivity in experience goods markets," *Quantitative Marketing and Economics*, 6, 139–176.

Erdem, T. and B. Sun (2002): "An Empirical Investigation of the Spillover Effects of Advertising and Sales Promotions in Umbrella Branding," *Journal of Marketing Research*, 39, 408–420.

Frank, R. E. (1962): "Brand Choice as a Probability Process," *Journal of Business*, 35, 43–56.

Fugger, N., E. Katok, and A. Wambach (2013): "Collusion in dynamic buyer-determined reverse auctions," Working Paper.

Gal-Or, E., M. Gal-Or, and A. Dukes (2007): "Optimal information revelation in procurement schemes," *RAND Journal of Economics*, 38, 400–418.

Gambardella, A., D. Harhoff, and B. Verspagen (2008): "The value of European patents," *European Management Review*, 5, 69–84.

Graham, S. J. H., B. H. Hall, D. Harhoff, and D. C. Mowery (2002): "Post-Issue Patent "Quality Control": A Comparative Study of US Patent Re-Examinations and European Patent Oppositions," U.C. Berkeley mimeo.

Hakenes, H. and M. Peitz (2008): "Umbrella branding and the provision of quality," *International Journal of Industrial Organization*, 26, 546–556.

Harhoff, D. and M. Reitzig (2004): "Determinants of opposition against EPO patent grants - the case of biotechnology and pharmaceuticals," *International Journal of Industrial Organization*, 20, 443–480.

Harhoff, D. and S. Wagner (2009): "The Duration of Patent Examination at the European Patent Office," *Management Science*, 55, 1969–1984.

Haruvy, E. and E. Katok (2013): "Increasing revenue by decreasing information in procurement auctions," *Production and Operations Management*, 22, 19–35.

Jap, S. D. (2002): "Online reverse auctions: Issues, themes and prospects for the future," *Journal of the Academy of Marketing Science*, 30, 506–525.

——— (2003): "An exploratory study of the introduction of online reverse auctions," *Journal of Marketing*, 67, 96–107.

Jap, S. D. and E. Haruvy (2008): "Interorganizational relationships and bidding behavior in industrial online reverse auctions," *Journal of Marketing Research*, 45, 550–561.

Jofre-Bonet, M. and M. Pesendorfer (2000): "Bidding behavior in a repeated procurement auction: A summary," *European Economic Review*, 44, 1006–1020.

Keane, M. P. (1997): "Modeling Heterogeneity and State Dependence in Consumer Choice Behavior," *Journal of Business and Economic Statistics*, 15, 310–326.

Klemperer, P. (1995): "Competition when Consumers have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade," *Review of Economic Studies*, 62, 515–539.

Kostamis, D., D. R. Beil, and I. Duenyas (2009): "Total-cost procurement auctions: Impact of suppliers' cost adjustments on auction format choice," *Management Science*, 55, 1985–1999.

Levin, J. and R. Levin (2002): "Patent Oppositions," Working paper.

Lewis, G. and P. Bajari (2011): "Procurement contracting with time incentives: Theory and evidence," *The Quarterly Journal of Economics*, 126, 1173–1211.

Massy, W. F. (1966): "Order and Homogeneity of Family Specific Brand-Switching Processes," *Journal of Marketing Research*, 3, 48–54.

Mizuno, T. (2003): "On the existence of a unique price equilibrium for models of product differentiation," *International Journal of Industrial Organization*, 21, 761–793.

Newton, M. A. and A. E. Raftery (1994): "Approximate Bayesian Inference with the Weighted Likelihood Bootstrap," *Journal of the Royal Statistical Society. Series B (Methodological)*, 56, 3–48.

Parkes, D. and J. Kalagnanam (2005): "Models of Iterative Multi-Attribute Vickrey Auctions," *Management Science*, 51, 435–451.

Pavlidis, P. and P. B. Ellickson (2012): "Switching Costs and Market Power Under Umbrella Branding," Simon School Working Paper.

Rezende, L. (2009): "Biased procurement auctions," *Economic Theory*, 38, 169–185.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2005): *Bayesian Statistics and Marketing*, New York: John Wiley & Sons, 1st ed.

Seetharaman, P. B., A. Ainslie, and P. K. Chintagunta (1999): "Investigating Household State Dependence Effects across Categories," *Journal of Marketing Research*, 36, 488–500.

Small, K. A. and H. S. Rosen (1981): "Welfare Economics with Discrete Choice Models," *Econometrica*, 49, 105–130.

Sobel, M. and W. Wei (2010): "Myopic Solutions of Homogeneous Sequential Decision Processes," *Operations Research*, 58, 1235–1246.

Stigler, G. J. and G. S. Becker (1977): "De Gustibus Non Est Disputandum," *The American Economic Review*, 67, 76–90.

Sullivan, M. (1990): "Measuring Image Spillovers in Umbrella-Branded Products," *The Journal of Business*, 63, 309–329.

Wan, Z. and D. R. Beil (2012): "RFQ auctions with supplier qualification screening," *Operations Research*, 57, 934–949.

Wan, Z., D. R. Beil, and E. Katok (2012): "When does it pay to delay supplier qualification? Theory and experiments," *Management Science*, 58, 2057–2075.

Weizsaecker, G. (2003): "Ignoring the rationality of others: evidence from experimental normal-form games," *Games and Economic Behavior*, 44, 145–171.

Wernerfelt, B. (1988): "Umbrella Branding as a Signal of New Product Quality: An Example of Signaling by Posting a Bond," *The RAND Journal of Economics*, 19, 458–466.