Knowing me, imagining you: Projection and overbidding in auctions

Breitmoser, Yves

22 January 2016

Knowing me, imagining you: Projection and overbidding in auctions

Yves Breitmoser*
Humboldt University Berlin

January 22, 2016

Abstract

Overbidding in auctions has been attributed to risk aversion, loser regret, level-k, and cursedness, though relying on different identifying assumptions. I argue that "type projection" organizes these findings and better captures observed behavior. Type projection formally models that people tend to believe others have object values similar to their own—a robust psychological phenomenon that naturally applies to auctions. First, I show that type projection implies the main behavioral phenomena in auctions, including increased sense of competition (like loser regret) and broken Bayesian updating (like cursedness). Second, re-analyzing data from seven experiments, I show that type projection explains the stylized facts of behavior across private and common value auctions. Third, in a structural analysis nesting existing approaches and emphasizing robustness, type projection consistently captures behavior best, in-sample and out-of-sample. The results reconcile bidding patterns across conditions and have implications for behavioral and empirical analyses as well as policy.

JEL–Codes: C72, C91, D44
Keywords: auctions, overbidding, winner's curse, projection, risk aversion, cursed equilibrium, limited depth of reasoning

*I thank Friedel Bolle, Jana Friedrichsen, Paul Heidhues, Matthias Lang, Theo Offerman, Sebastian Schweighofer-Kodritsch, Jean-Robert Tyran, Felix Weinhardt, Georg Weizsäcker, seminar audiences in Berlin, Nottingham and Regensburg, and conference participants in Montreal (ESWC 2015) and Münster (VfS Jahrestagung, 2015) for helpful comments. Financial support of the DFG (project BR 4648/1–1) is greatly appreciated. Address: Spandauer Str. 1, 10099 Berlin, Germany, email: yves.breitmoser@hu-berlin.de, Telephone/Fax: +49 30 2093 99408/5619.

1 Introduction

The false consensus bias is the tendency to assume that one's own opinions, preferences and values are typical and shared by others. Following Ross et al. (1977), such "projection" has been confirmed in many experiments (Mullen et al., 1985). Projection may persist even if subjects are provided with factually contradicting information (Krueger and Clement, 1994). Thus, projection is of intuitive relevance in all choices under incomplete information—not just in the non-strategic environments on which the psychological literature traditionally focuses, but also in strategic interactions. Current concepts studying projection in "games" focus on one-sided incomplete information. In their seminal paper, Loewenstein et al. (2003) study projection of utility onto future selves, finding that it explains anomalies in purchases of durable goods. In a different context, Madarász (2012) studies projection of information from an informed player to an uninformed one, which explains the hindsight bias in agency problems.

In the present paper, I argue that projection of "types" appears to affect behavior in auctions. Auctions are widely discussed games with two-sided incomplete information, and types capture signals about object values. I introduce a simple model of type projection where players may overestimate the probability that their opponents share their type—ranging from zero projection (the original Bayesian case) to full projection (disregarding all prior information).[1] The degree of projection is denoted by $\rho \in [0, 1]$. An equilibrium exists for each ρ, and in equilibrium, players anticipate their opponent types' actual strategies but compute their expected payoffs projecting their own type.

The observations supporting type projection, and some implications, are summarized as follows. Type projection formalizes a phenomenon observed by a large literature of psychological work, applies naturally in auctions, and generates the behavioral phenomena observed across conditions: loser regret in first-price auctions and broken Bayesian updating in common value auctions. It captures stylized facts more comprehensively than existing concepts and fits behavior very robustly, across information conditions, across experiments, and across identifying assumptions (on strategic beliefs). The degree of projection ρ is largely invariant, around 0.5, explaining the robustness and corroborating that projection is of first-order relevance. The results have policy implications, as the projection bias is reduced when subjects are educated explicitly (Engelmann and Strobel, 2012), which enables efficiency gains, and implications for behavioral and empirical work. For, type projection intuitively factors in all symmetric Bayesian games (e.g. on social preferences under anonymity, see also Blanco et al., 2014), and in empirical work, as projection fits robustly across private and common values, thus relating to empirical auctions which tend to be hybrid (Haile, 2001; Goeree and Offerman, 2002).

The underlying intuition also is simple. Type-projecting bidders project their signals or values. Values are known to be projected in general, e.g. in bargaining (Bottom and Paese, 1999; Galinsky and Mussweiler, 2001) and in consumption decisions (Frederick, 2012; Kurt and Inman, 2013). As for auctions, consider bidding to buy a house. Projecting bidders neglect competitors whose values are vastly inferior, against whom they surely win, and competitors whose values are vastly superior, against whom they surely

[1] Full projection is regularly considered in analyses of social preferences. The present paper considers the more intricate case of imperfect projection. Allowing for imperfect projection is critical, as full projection is neither observed in psychology nor fits bidding in auctions.

lose. They focus on competitors with similar values, trying to ensure winning against them. This focus increases the sense of competition and obscures the perceived value distribution, which impacts bidding as follows.

In private value auctions, projecting bidders overbid, as they overestimate the share of opponents with similar values and outbid them to increase the probability of winning. In contrast, risk aversion emphasizes a trade-off between increasing winning probability and increasing conditional profit. Following Engelbrecht-Wiggans (1989), the former relates to loser regret (regret of losers if they could have won profitably) and the latter relates to winner regret (regret of winners if they could have won with lower bids). Filiz-Ozbay and Ozbay (2007) find that subjects do not trade off these regrets but focus on loser regret. This focus contradicts risk aversion and is implied by type projection. In addition, I find that subjects randomize consistently and use left-skewed mixed strategies. Again, this contradicts risk aversion and supports type projection, which predicts mixed equilibria with, indeed, left-skewed strategies.

In common value auctions, type projection obscures the perceived type distribution and thus weakens Bayesian updating, similarly to cursed equilibrium (Eyster and Rabin, 2005). Cursed bidders believe their opponents get random signals with probability χ (and the true signals with $1-\chi$), and type projecting bidders believe their opponents have signals similar to their own with probability ρ (and the true signals with probability $1-\rho$). In both cases, bidders underestimate the informativeness of their opponents' bids and experience the Winner's Curse. The reason for broken Bayesian updating is very similar, and both concepts are compatible with the intuition usually expressed by economists (e.g. Milgrom, 1989), but between the two approaches, the belief perturbation underlying type projection is supported by independent psychological evidence. This evidence (on false consensus) addresses interactions with symmetric type sets, and in turn, cursed equilibrium appears more appropriate to model games with asymmetric type sets (e.g. buyer-seller interactions). In addition, subjects randomize consistently even in common value auctions, which supports the mixed equilibrium predicted by type projection. Subjects overbid with both private and common values, but more so under common values, which also confirms the prediction of type projection.

Relatedly, let me discuss the notion that jointly, risk aversion and cursedness explain bidding across conditions. I argue that, even jointly, these explanations are inconsistent with the received intuition: Bidders experience loser regret in first-price auctions (Filiz-Ozbay and Ozbay, 2007) and broken Bayesian updating of common object values (Milgrom, 1989). Focusing on first-price auctions, an explanation consistent with the received intuition induces an incentive similar to loser regret across conditions, and mistaken Bayesian updating on top of it in common value auctions. This is particularly intuitive if one compares auctions with affiliated private values and auctions with common value, as their differences are very small.[2] The model combining risk aversion and cursed equilibrium is not consistent in this sense, as an incentive similar to loser regret is implied only in private value auctions—risk aversion is outcome irrelevant in (standard) common value auctions.

[2] Detailed definitions follow, but briefly, let X_0 denote a random variable revealed only to Nature, and let (X_i) denote the individual signals revealed privately to the bidders (i). All X_i are distributed (i.i.d.) on an interval around X_0. The sole difference between common values and affiliated private values is that player i's object value is $v = X_0$ (common value) or $v = X_i$ (private value).

Type projection is consistent also in this sense. It induces an incentive to "avoid loser regret" in any first-price auction and breaks Bayesian updating in common value auctions. Thus, to summarize, type projection formalizes a robust psychological finding, is intuitively consistent across conditions and fits a wide range of stylized facts. These results are complemented by a structural analysis of bidding on a large data set comprising seven experiments. The data set forms the union of the data sets analyzed in seminal analyses of bidding behavior, which limits selection effects in favor of type projection. In addition, merging multiple data sets allows me to assess whether models are precise (in-sample) and reliable (out-of-sample). Both features are desirable in behavioral and empirical analysis, but reliability will be of particular relevance here.

To clarify why, let me briefly review existing results. Goeree et al. (2002b) and Bajari and Hortacsu (2005) show that risk aversion captures bidding in private value auctions, Filiz-Ozbay and Ozbay (2007) and Engelbrecht-Wiggans and Katok (2007) observe loser regret, Eyster and Rabin (2005) observe cursedness in common value auctions, and Crawford and Iriberri (2007) observe limited depth of reasoning in either condition. That is, the results vary enormously between studies. The main reason appears to relate to the identifying assumptions imposed on strategic beliefs, which range from naive beliefs (level-1) over Nash beliefs (equilibrium without anticipating errors) to rational expectations. To reconcile these results, specific and extreme assumptions on belief formation are to be avoided. I introduce a concept based on quantal response equilibrium (McKelvey and Palfrey, 1995) that nests the three belief models above and endogenizes the assumption on belief formation. While this solves one problem, Haile et al. (2008) suspect that generalized forms of QRE may overfit and lack robustness themselves. The data used here allow me to directly address this issue by evaluating robustness, i.e. the accuracy of predictions across experiments.[3] In addition, this analysis verifies whether the models are applicable across data sets, e.g. in (future) analyses of different data.

The results of this analysis confirm the qualitative evidence outlined above. Type projection indeed captures behavior best, both descriptively (in-sample) and predictively (out-of-sample). Further, inexperienced subjects tend to underestimate the rationality of others, though not in the way predicted by level-k. As subjects gain experience, their beliefs approach rational expectations, the precision in maximizing utility increases, subject heterogeneity becomes significant, but the degree of projection remains largely constant (around 0.5). Type projection is comprehensive in the sense that neither risk aversion nor cursedness capture facets of behavior incompatible with projection, and the results are robust to "non-standard" information conditions. Thus, I conclude that the book on behavior in auctions, which appeared closed for a while, may have to be reopened.

Section 2 introduces the model of type projection, Section 3 introduces the data sets. Sections 4 and 5 analyze the relations to stylized facts. Sections 6 and 7 contain the structural analysis of bidding. Results and implications are discussed in Section 8.

[3] Another issue with using the generalization of QRE is that the underlying QRE needs to be computed explicitly—the fixed point computation cannot be avoided using the insight of Bajari and Hortacsu (2005), by exploiting rational expectations, as relaxing rational expectations is exactly the point. The explicit computation of QREs is computationally demanding in standard auctions, due to the complexity of randomized bidding functions, but a novel observation allows me to reduce the strategy complexity by an order of magnitude and thus enables computation of QREs using massive parallelization (on GPUs).

2 Behavioral models of bidding

In this section, I define type projection and discuss its relation to existing concepts. The necessary notation is standard. $N = \{1, \ldots, n\}$ denotes the set of players, A_i and T_i, $i \in N$, denote the sets of actions and types, respectively. The set of Nature's types is T_0, and the set of type profiles is $T = T_0 \times T_1 \times \cdots \times T_n$. Type and action sets are finite, as e.g. in all laboratory auctions. The prior probability that the type profile is $t \in T$ is $\Pr(t)$ and the prior that i's type is t_i is $\Pr_i(t_i)$. A game is called *type-symmetric* if $T_i = T_j$ and $\Pr_i = \Pr_j$ for all $i, j \in N$. Action profiles are denoted by $a \in A = A_1 \times \cdots \times A_n$. Given action profile $a \in A$ and type profile $t \in T$, i's payoff is $p_i(a,t)$. As usual, $A_{-i} = \bigtimes_{j \in N \setminus \{i\}} A_j$ and $T_{-i} = \bigtimes_{j \in N \cup \{0\} \setminus \{i\}} T_j$, and the posterior of $t_{-i} \in T_{-i}$ given t_i is $\Pr(t_{-i}|t_i)$.

The strategy $\sigma_i(\cdot|t_i) \in \Delta A_i$ of i maps i's actions to probabilities contingent on type t_i. The expected payoff of type $t_i \in T_i$ from action a_i in response to σ_{-i} is

$$\pi_i\big(a_i|t_i,\sigma_{-i}\big) = \sum_{t_{-i} \in T_{-i}} \sum_{a_{-i} \in A_{-i}} \Pr(t_{-i}|t_i)\ p_i\big[(a_i,a_{-i}),(t_i,t_{-i})\big] \prod_{j \neq i} \sigma_j(a_j|t_j). \tag{1}$$

Given payoffs π_i, the set of best responses of type t_i to σ_{-i} is

$$BR_{t_i}(\sigma_{-i}|\pi_i) = \arg \max_{\sigma_i' \in \Delta A_i} \sum_{a_i \in A_i} \sigma_i'(a_i)\ \pi_i\big(a_i|t_i,\sigma_{-i}\big). \tag{2}$$

A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a Bayesian Nash equilibrium (BNE) if all types $t_i \in T_i$ of all players $i \in N$ choose best responses, $\sigma_i(\cdot|t_i) \in BR_{t_i}(\sigma_{-i}|\pi_i)$.

2.1 Type projection

Type projection equilibrium (TPE) extends Bayesian Nash equilibrium by incorporating the projection bias. This can be modeled in a variety of ways, but to prevent over-representation of the adequacy of type projection, I set up a simple model not exploiting degrees of freedom. Generalizations and alterations are discussed in Breitmoser (2015).

Type-projecting players assign weight $1-\rho$, $\rho \in [0,1)$, to the objective prior Pr and weight ρ to their projection that all opponents' types are equal to their type. The parameter ρ is called *degree of projection*, the original Bayesian belief obtains for $\rho = 0$. A projecting t_i thus expects that action a_i yields the *projected payoff*

$$\begin{aligned}\pi_i^{Proj}\big(a_i|t_i,\sigma_{-i}\big) = (1-\rho) &\sum_{t_{-i} \in T_{-i}} \sum_{a_{-i} \in A_{-i}} \Pr(t_{-i}|t_i)\ p_i\big[(a_i,a_{-i}),(t_i,t_{-i})\big] \prod_{j \neq i} \sigma_j(a_j|t_j) \\ +\rho &\sum_{t_{-i} \in T_{-i}} \sum_{a_{-i} \in A_{-i}} \Pr(t_{-i}|t_i)\ p_i\big[(a_i,a_{-i}),(t_i,t_{-i})\big] \prod_{j \neq i} \sigma_j(a_j|t_i).\end{aligned} \tag{3}$$

For any $\rho \in [0,1)$, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a ρ-type projection equilibrium (ρ-TPE) if all types $t_i \in T_i$ of all players $i \in N$ choose best responses under projection, $\sigma_i(\cdot|t_i) \in BR_{t_i}(\sigma_{-i}|\pi_i^{Proj})$. In type-symmetric games, existence obtains as any ρ-TPE of a Bayesian game Γ is a BNE of an augmented game $\tilde{\Gamma}$ where the projected events are possible draws by Nature (Breitmoser, 2015). This relates to the argument establishing existence of cursed equilibrium in Eyster and Rabin (2005).

The above formulation of projection is simple in that assumes that players project their types onto *all* their opponents simultaneously and in that players project their *exact* type. The correlated projection facilitates tractability and reflects the observations of Camerer et al. (2004) and Costa-Gomes et al. (2009) that subjects believe their opponents make correlated choices. An alternative assumption would be that projection is independent across opponents, but I am not aware of independent evidence supporting it. An alternative to the latter assumption would be "fuzzy" projection, i.e. the belief that opponents' types are "similar" to the own type. This would consume additional degrees of freedom (e.g. a distance measure for types and a function mapping type distance to degree of projection) which I seek to prevent for transparency.

Type projection differs distinctly from other forms of projection discussed in the literature, e.g. utility projection and information projection. Loewenstein et al. (2003) introduce utility projection: Given consumption c and current state s, the decision maker predicts the utility will be $(1-\alpha)\,u(c,s')+\alpha u(c,s)$, $\alpha\in[0,1]$, in future state s'. Applying the idea to multi-player games, utility projection implies that players believe their opponents' types have "average" utilities and play pure strategies each. In contrast, a type-projecting player associates each list of opponents' types t_{-i} with *mixed strategies*. With probability $1-\rho$ the true types t_{-i} play and with probability ρ the projected types play.[4] In turn, Madarász (2012) introduces information projection, i.e. a player believing his opponents know all he knows, in addition to their existing knowledge. In auctions, information projection implies that the opponents know his value, in addition to knowing their own values. Type projection implies that the opponents share i's value. Information projection is appealing in cases of one-sided "missing" information, and it provides an intriguing explanation of the hindsight bias, but it appears less appealing in auctions—where there is no objectively "missing" information, but heterogeneity of types.[5]

2.2 Alternative concepts

Cursed equilibrium The relation to cursed equilibrium (Eyster and Rabin, 2005) has been discussed in the Introduction: Both concepts assume that players have a mistaken understanding of the type distribution. Given the degree of cursedness $\chi\in[0,1]$, cursed players assign weight $1-\chi$ to the Bayesian case and χ to the event that their opponents' types are random and uninformative given the own private information. In the latter case, the opponents play the average strategy $\overline{\sigma}_j(a_{-i}|t_i)=\sum_{t_{-i}\in T_{-i}}\Pr(t_{-i}|t_i)\prod_{j\neq i}\sigma_j(a_j|t_j)$, and overall, cursed players expect payoffs

$$\begin{aligned}\pi_i^{Curse}\big(a_i|t_i,\sigma_{-i}\big)=(1-\chi)&\sum_{t_{-i}\in T_{-i}}\sum_{a_{-i}\in A_{-i}}\Pr(t_{-i}|t_i)\;p_i\big[(a_i,a_{-i}),(t_i,t_{-i})\big]\prod_{j\neq i}\sigma_j(a_j|t_j)\\&+\chi\sum_{t_{-i}\in T_{-i}}\sum_{a_{-i}\in A_{-i}}\Pr(t_{-i}|t_i)\;p_i\big[(a_i,a_{-i}),(t_i,t_{-i})\big]\,\overline{\sigma}_j(a_{-i}|t_i).\end{aligned}\tag{4}$$

[4] Also note the difference to "strategy" projection: A type projecting player believes opponents share his type but keep their individual incentives. Both utility projecting and strategy projecting players implicitly assume the opponents neglect their original incentives and adopt his utilities or strategies.

[5] Madarász (2015) generalizes the concept by including "ignorance projection" and applies it to games with two-sided incomplete information. The differences still appear major, as information projection appears to predict pure equilibria in auctions (since payoffs are continuous), but precise comparisons are impossible, as the shape of equilibrium strategies under information projection in auctions (which are not the main application) is not characterized (see Example 2.1.2 in Madarász, 2015).

A strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a χ-cursed equilibrium if $\sigma_i(\cdot|t_i) \in BR_{t_i}(\sigma_{-i}|\pi_i^{Curse})$ for all i and t_i. I am not aware of independent evidence supporting "random projection" as in cursed equilibrium (as opposed to projection of the own type), but cursed equilibrium appears well-suited to capture beliefs if type sets are asymmetric. Market interactions with one-sided incomplete information as analyzed in Eyster and Rabin (2005) are a prototypical example. In such asymmetric games, type projection appears less intuitive.

Risk aversion Cox et al. (1985, 1988) argue that a potential factor in bidding is constant relative risk aversion (CRRA), $u(p) = p^\alpha/\alpha$ with $\alpha \neq 0$, with expected utilities

$$\pi_i^{CRRA}\left(a_i|t_i,\sigma_{-i}\right) = \sum_{t_{-i}\in T_{-i}} \sum_{a_{-i}\in A_{-i}} \Pr(t_{-i}|t_i)\; u\big(p_i\left[(a_i,a_{-i}),(t_i,t_{-i})\right]\big) \prod_{j\neq i} \sigma_j(a_j|t_j). \quad (5)$$

CRRA utilities $u(\cdot)$ can equally be used to complement projection and cursedness. As it stands, risk aversion is the leading explanation of overbidding in private value auctions, but the more recent observations on loser regret, e.g. Filiz-Ozbay and Ozbay (2007) and Engelbrecht-Wiggans and Katok (2007), challenge this perspective (as discussed above).

Limited depth of reasoning The concepts discussed so far have in common that they are defined in terms of the payoff structure $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$. The players' beliefs about their opponents' strategies are taken as given. The complementary approach is to vary the belief system, allowing that players deviate from BNE by violating rational expectations.[6] The seminal model in this strand literature, level-k, follows Stahl and Wilson (1995) and Nagel (1995); other belief systems are discussed below. Assuming level-0 randomizes uniformly, $\sigma^0(\cdot|t_i) = 1/|A_i|$ for i, t_i, and given a payoff structure $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$, player i has level-k depth of reasoning, $k \geq 1$, if he plays $\sigma^k(\cdot|t_i) \in BR_{t_i}(\sigma_{-i}^{k-1}|\tilde{\pi}_i)$ for all t_i. In a similar manner, level-k has been applied to auctions by Crawford and Iriberri (2007).

3 Data sources

I am analyzing data sets from multiple experiments to address the risk of misinterpreting model adequacy. Broadening the data basis reduces the fallacy to overfitting, and the joint analysis of multiple data sets allows me to assess predictive adequacy across experiments: Fit parameters to some set of experiments and predict behavior in the remaining experiments. Evaluating predictive adequacy across experiments addresses overfitting most transparently by showing to which degree the results on payoff structure and belief system obtained here will be helpful in (future) analyses of different data sets. Finally, pooling auctions under varying information conditions, independent private values, affiliated private values, and common values, allows us to examine robustness to real-world conditions, as they tend to be hybrid (Haile, 2001; Goeree and Offerman, 2002).

[6] Note that both cursedness and projection can equally be defined as concepts relaxing the belief structure. Above, they have been defined in terms of the payoff structure, as both Eyster and Rabin (2005) and the above definitions emphasize that an equilibrium assumption is maintained even under cursedness and projection (a BNE of an augmented game), while standard models of alternative belief systems (such as level-k) emphasize the non-equilibrium character of the predicted strategy profiles.

Table 1: Data sources

Format	Source	Values	Signals	Inexperienced #Subj	Inexperienced #Obs	Experienced #Subj	Experienced #Obs
Standard auctions							
First price,	Kagel and Levin (2002)	$v = X_0$	$X_i \| X_0 \sim U[s \pm \varepsilon]$	51	255		
common value	Kagel and Levin (1986)	$v = X_0$	$X_i \| X_0 \sim U[s \pm \varepsilon]$			49	237
Second price, common value	Garvin and Kagel (1994)	$v = X_0$	$X_i \| X_0 \sim U[s \pm \varepsilon]$	28	140		
First price, affiliated private	Kagel et al. (1987)	$v = X_i$	$X_i \| X_0 \sim U[x_0 \pm \varepsilon]$	42	210	42	210
First price,	Dyer et al. (1989)	$v = X_i$	$X_i \sim U[0, 30]$	18	180	18	180
Independ. private	Kagel and Levin (1993)	$v = X_i$	$X_i \sim U[0, 28.3]$	10	50	10	100
Non-standard auctions							
First price, Independ. private	Goeree et al. (2002b)	$v = X_i$	X_i discrete	80	400	80	400
Second price, Common value	Avery and Kagel (1997)	$v = X_1 + X_2$	$X_i \sim U[1, 4]$	23	115	23	115

Note: The discrete signals in Goeree et al. (2002b) are uniform draws from either $\{0,2,4,6,8,11\}$ or $\{0,3,5,7,9,12\}$. The data for inexperienced subjects are mostly from Crawford and Iriberri (2007). In most rounds of Dyer et al. (1989) and Kagel and Levin (1993), the subjects played two auction markets simultaneously. Focusing on the first and last five rounds they played, we mostly have ten observations per subject. Due to bankruptcies in CV auctions, there are not always five observations per subject.

The data sources are listed in Table 1. These seven data sets form the union of the data sets analyzed in the most influential studies of bidding behavior, Goeree et al. (2002b), Bajari and Hortacsu (2005), Eyster and Rabin (2005), and Crawford and Iriberri (2007). The repetitive re-analysis of these data sets indicates consensus on their adequacy to study bidding behavior, and re-analyzing these very data sets implies that if data selection influences the results, it would be in favor of existing theories.

I distinguish "standard" auctions and "non-standard" auctions. An auction is labeled "standard" if signals and bids are (approximately) continuous and signals are distributed independently conditional on the own object value. Kagel and Levin (1986) and Kagel and Levin (2002, Chapter 4) analyze first-price, common-value auctions. The common value is $v = X_0$ and individual signals are distributed as $X_i|X_0 \sim U[x_0 \pm w]$. The BNE strategy is $b(x_i) \approx x_i - w$.[7] Kagel and Levin (1986) and Garvin and Kagel (1994) analyze second price, common value auctions. Signals and value are as in the first-price case, but the BNE strategy is $b(x_i) = x_i - w + \frac{2w}{n}$, with n as the number of players. Kagel et al. (1987) analyze first-price auctions with affiliated private values. The private value $v = X_i$ is distributed as $X_i|X_0 \sim U[x_0 \pm w]$ with BNE strategy $b(x_i) \approx x_i - \frac{2w}{n}$. Finally, Dyer et al. (1989) and Kagel and Levin (1993) analyze first-price auctions with independent private values. The private value $v = X_i$ is distributed as $X_i \sim U[0, 30]$ and $X_i \sim U[0, 28.3]$, respectively. The BNE strategy is $b(x_i) = x_i(n-1)/n$. The "non-standard" auctions are reviewed below, as they will be used primarily in out-of-sample tests of the models (they appear particularly challenging for the intuition underlying type projection).

[7] The exact BNE strategy is $b(x_i) = x_i - w + Y$ with $Y = \frac{2w}{n+1} \times \exp\{-n(x_i - \underline{x} - w)/2w\}$, but $Y \approx 0$ if the signal x_i is not very close to the bounds of the signal space.

Finally, I comparatively analyze experienced subjects and inexperienced subjects. This comparison slightly extends the literature, which analyzes either inexperienced subjects (Crawford and Iriberri, 2007) or experienced ones (most other studies). In particular, depth of reasoning and rationality of expectations are argued to vary with experience (e.g. Crawford and Iriberri, 2007): initial behavior (inexperienced subjects) is intuitively closer to level-k and converged behavior (experienced subjects) is intuitively closer to rational expectations and equilibrium. Following Crawford and Iriberri (2007), a subject is called "inexperienced" during the first five auctions, and by inversion, "experienced" during the last five auctions (out of some 20 auctions in a session).[8]

4 Bidding in auctions: Stylized facts

This section examines the basic characteristics of bidding, including complexity and heterogeneity, which are critical for the subsequent econometric models, and the first three moments of the bid distributions. This will allow us to compare the qualitative predictions of the models. Throughout the paper, I report the results of statistical tests at two levels of significance: 0.05 and 0.005. The former is standard, and the latter implements the Bonferroni correction assuming 10 simultaneous tests across treatments and models, which will generally be adequate per level of experience.

4.1 Complexity of bidding functions

How do bids relate to signals? In first-price auctions, equilibrium bids are smaller than conditional object values, i.e. subjects shade bids to ensure positive profits. In the standard auctions in Table 1, BNE predicts bid shading by either absolute or relative reductions—relative reductions for independent private values (IPV) and absolute reductions for affiliated private values (APV) and common values (CV). I test these predictions, as the complexity of the bidding functions affects all future steps of the analysis. The alternative hypothesis is that subjects do not strictly adhere to either, relative reductions in IPV auctions and absolute reductions in APV and CV auctions. Instead, subjects may engage in mixtures of relative and absolute reductions, e.g. high relative reductions if they have high signals and low relative reductions if they have low signals.

Strategy complexity can be tested by estimating both intercept and slope of the bidding functions $b(x)$. In IPV auctions, the intercept is predicted to be zero, in APV and CV auctions the slope is predicted to be one.[9] Specifically, in IPV auctions, I estimate $b = \alpha + \beta \cdot x$, and the BNE prediction is $\alpha = 0$ and $\beta < 1$. In APV and CV auctions, I estimate $b = \alpha \cdot w + \beta \cdot x$, including the "signal bandwidth" w which is a constant within treatments—thus, α represents the intercept, but controlling for w facilitates comparisons

[8] In common value auctions, in particular, behavior has not converged after five auctions, which precludes me from using all observations from the sixth auction on in the analysis of experienced subjects. In turn, behavior is independent of time during the first five auctions and during the last five auctions, respectively (as shown in the supplementary material), indicating that these partitions of the data set meet the time invariance assumed in the analysis.

[9] As reviewed above, the equilibrium bids are $b(x_i) \approx x_i - w$ and $b(x_i) \approx x_i - \frac{2w}{N}$ in the "standard" CV and APV auctions, respectively, and they are $b(x) = n/(n-1) \cdot x$ in the IPV auctions.

Table 2: Summary statistics of bidding in the first-price auctions

	Inexperienced subjects					Experienced subjects				
		Degree of	Standard Deviation				Degree of	Standard Deviation		
Condition	Bidding function	Overbidding	within Ss	between Ss	Skewness	Bidding function	Overbidding	within Ss	between Ss	Skewness
Independent private values, First price (DKL89, KL93)										
$N=3$	$b=\underset{(0.234)}{-0.031}+\underset{(0.018)}{0.803^{\star\star}}\cdot x$	$\underset{(0.017)}{0.104^{\star\star}}$	$\underset{(0.031)}{0.161}$	0.025	$-3.17^{\star\star}$	$b=\underset{(0.149)}{0.028}+\underset{(0.01)}{0.822^{\star\star}}\cdot x$	$\underset{(0.014)}{0.143^{\star\star}}$	$\underset{(0.018)}{0.126}$	0.054	$-3.26^{\star\star}$
$N=6$	$b=\underset{(0.196)}{-0.039}+\underset{(0.013)}{0.849^{\star\star}}\cdot x$	$\underset{(0.02)}{-0.021}$	$\underset{(0.034)}{0.162}$	0.053	$-3.35^{\star\star}$	$b=\underset{(0.182)}{0.037}+\underset{(0.009)}{0.875^{\star\star}}\cdot x$	$\underset{(0.012)}{0.034^{\star\star}}$	$\underset{(0.018)}{0.108}$	0.044	$-4.49^{\star\star}$
$N=5$	$b=\underset{(0.241)}{0.195}+\underset{(0.01)}{0.886^{\star\star}}\cdot x$	$\underset{(0.021)}{0.08^{\star\star}}$	$\underset{(0.053)}{0.145}$	0.028	$-4.34^{\star\star}$	$b=\underset{(0.496)}{-0.873}+\underset{(0.017)}{0.896^{\star\star}}\cdot x$	$\underset{(0.042)}{-0.021}$	$\underset{(0.036)}{0.264}$	0.129	$-1.65^{\star}$
Affiliated private values, First price (KHL87)										
$N=6, w=6$	$b=\underset{(0.006)}{0.986^{\star}}\cdot x-\underset{(0.079)}{0.284^{\star\star}}\cdot w$	$\underset{(0.038)}{-0.127^{\star\star}}$	$\underset{(0.085)}{0.366}$	0.148	$-3.83^{\star\star}$					
$N=6, w=12$	$b=\underset{(0.006)}{0.992}\cdot x-\underset{(0.048)}{0.172^{\star\star}}\cdot w$	$\underset{(0.026)}{0.104^{\star\star}}$	$\underset{(0.006)}{0.052}$	0.15	0.1	$b=\underset{(0.005)}{1}\cdot x-\underset{(0.037)}{0.247^{\star\star}}\cdot w$	$\underset{(0.021)}{0.088^{\star\star}}$	$\underset{(0.044)}{0.168}$	0.129	$-1.52^{\star}$
$N=6, w=24$						$b=\underset{(0.006)}{1}\cdot x-\underset{(0.02)}{0.168^{\star\star}}\cdot w$	$\underset{(0.022)}{0.164^{\star\star}}$	$\underset{(0.015)}{0.09}$	0.135	-0.24
Common value auctions, First price (KL86)										
$N\le 4, w=6$	$b=\underset{(0.003)}{0.996}\cdot x-\underset{(0.139)}{0.22}\cdot w$	$\underset{(0.066)}{0.657^{\star\star}}$	$\underset{(0.033)}{0.341}$	0.28	$0.57^{\star\star}$					
$N\le 4, w\le 18$	$b=\underset{(0.013)}{1.014}\cdot x-\underset{(0.172)}{0.676^{\star\star}}\cdot w$	$\underset{(0.099)}{0.551^{\star\star}}$	$\underset{(0.163)}{0.43}$	0.095	$0.54^{\star}$	$b=\underset{(0.016)}{1.002}\cdot x-\underset{(0.576)}{0.905^{\star\star}}\cdot w$	$\underset{(0.038)}{0.106^{\star}}$	$\underset{(0.087)}{0.276}$	0.011	-1.77
$N\le 4, w\ge 24$						$b=\underset{(0.021)}{1}\cdot x-\underset{(0.091)}{0.63^{\star\star}}\cdot w$	$\underset{(0.05)}{0.373^{\star\star}}$	$\underset{(0.058)}{0.31}$	0.178	$0.7^{\star}$
$N=7, w=6$	$b=\underset{(0.002)}{0.999}\cdot x-\underset{(0.088)}{0.322^{\star\star}}\cdot w$	$\underset{(0.067)}{0.629^{\star\star}}$	$\underset{(0.036)}{0.333}$	0.313	$0.54^{\star}$					
$N\ge 5, w=12$						$b=\underset{(0.007)}{0.99}\cdot x-\underset{(0.084)}{0.575^{\star\star}}\cdot w$	$\underset{(0.051)}{0.338^{\star\star}}$	$\underset{(0.025)}{0.151}$	0.225	$1.02^{\star}$
$N\ge 5, w=18$						$b=\underset{(0.008)}{1}\cdot x-\underset{(0.082)}{0.654^{\star\star}}\cdot w$	$\underset{(0.045)}{0.348^{\star\star}}$	$\underset{(0.025)}{0.296}$	0.206	$1.09^{\star\star}$
$N\ge 5, w\ge 24$						$b=\underset{(0.012)}{0.999}\cdot x-\underset{(0.085)}{0.714^{\star\star}}\cdot w$	$\underset{(0.046)}{0.279^{\star\star}}$	$\underset{(0.025)}{0.231}$	0.201	$1.33^{\star}$

Notation: b is the bid, x is the signal, w is the interval width in APV and CV auctions
Normalized bids: The normalized bids are $r=(b-x)/w$ in APV and CV auction and $r=b/x$ in IPV auctions
Degree of overbidding is the difference between the mean normalized bid and the normalized equilibrium bid (BNE without risk aversion), it is estimated controlling for subject-level random effects ("between-subject standard deviation"). The within- and between-subject standard deviations refer to the distribution of normalized bids
Skewness: Skewness of the normalized bids after controlling for subject-level random effects (i.e. skewness of the errors in the regressions of normalized bids on intercept).
Experience: Subjects are "inexperienced" in their first five auctions and "experienced" in their last five auctions.
Asterisks indicate the bootstrapped p-values (see Footnote 10) of the null hypotheses that the respective parameters are either 1 (in case of the coefficients of x in APV and CV auctions, which are predicted to be 1) or 0 (in all other cases). "$\star\star$" indicates p-values less than .005, and "$\star$" indicates p-values between .005 and .05. The lower threshold .005 implements the Bonferroni correction for multiple testing across treatments (for around 10 treatments per level of experience).

across treatments. The BNE prediction is $\alpha < 0$ and $\beta = 1$. I include subject-level random effects and bootstrap p-values to account for the panel structure of the data.[10]

Table 2, column "Estimate $B(X)$", presents the results. In APV and CV auctions, the estimate of β differs significantly from 1 in one of the twelve treatments (at $\alpha = .05$), which is well within the limits of chance. In IPV auctions, α is insignificantly different form zero in all cases, suggesting that subjects indeed make relative reductions. The estimated parameters are also economically insignificant, i.e. small in relation to the range of signals. Thus, I conclude that the theoretical prediction of either absolute or relative reductions is validated.

Result 1. *As predicted by BNE, in APV and CV auctions normalized bids $(b-x)/w$ are independent of x, and in IPV auctions normalized bids b/x are independent of x.*

Without loss of information, both theoretically and empirically, we may thus represent bidding functions by scalars capturing the respective relative or absolute reductions. I refer to these scalars as *normalized bids*. In APV and CV auctions, the normalized bid is $r = (b-x)/w$. and in IPV auctions, it is $r = b/x$. One may think of r as the inverted degree of bid shading.[11] Values close to 0 in APV and CV auctions, or close to 1 in IPV auctions, indicate zero bid shading.

Mixed strategies are distributions over normalized bids and thus one-dimensional. The one-dimensionality enables analyses of average behavior aggregated over signals and condensed plots of bid distributions in histograms (Figure 1). At least as importantly, the condensation to one dimension is computationally critical. To see this, consider the ex-ante hypothesis that mixed strategies are distributions over bidding functions, i.e. at least two-dimensional. As numerical example, in auctions with 100 possible bids for each of 100 possible signals, we need to consider merely one distribution over 100 normalized bids instead of 100 distributions over 100 bids. Computing a fixed point (a mixed equilibrium) in 100 probabilities is computationally feasible, but computing one in 100^2 probabilities is not (yet). Thus, the reduction of strategy complexity allows us to compute fixed points and apply novel belief models, namely concepts relaxing assumptions such as rational expectations or Nash beliefs. This is discussed in detail below.

4.2 Is subject heterogeneity continuous or discrete?

To construct efficient econometric models, we need to capture subject heterogeneity: Are subjects clustered, i.e. is heterogeneity captured only by a mixture of several (normal) distributions? The latter is predicted by level-k, and thus, testing multimodality also is a test of level-k. The histograms in Figures 1–3, and the respective kernel density estimates, suggest that the distributions are unimodal. To rigorously test multimodality, I estimate

[10] The bootstrap is implemented as follows. The data set is resampled $R = 10.000$ times at the subject level (reflecting the panel structure of the data). To define the p-value of the null hypothesis that some statistic s is zero, let s_b denote its value in sample b and let s_0 denote its original value. The p-value of the two-sided test is $\frac{1}{2R}\#\{b : |s_b - \bar{s}| > |s_0|\} + \frac{1}{2R}\#\{b : |s_b - \bar{s}| \geq |s_0|\}$, where $\bar{s}$ is the mean of (s_b) and R the number of samples. Other p-values are defined analogously.

[11] The lower the normalized bid, the higher the degree of bid shading. For example, in APV and CV auctions, with $r = -0.4$, subjects bid $0.4 \cdot w$ less than their signal, $r = 0$ indicates bidding one's signal, $r = -2/N$ is the BNE strategy in APV auctions, and $r = -1$ is the BNE strategy in CV auctions.

Figure 1: First-price auctions, affiliated private values (KHL87). Inexperienced subjects (a–b) vs. experienced subjects (c–d). Plots are histograms of $r = (Bid - Signal)/w$

(a) Inexp: $N = 6, w = 6$ (b) Inexp: $N = 6, w = 12$ (c) Exp: $N = 6, w = 12$ (d) Exp: $N = 6, w = 24$

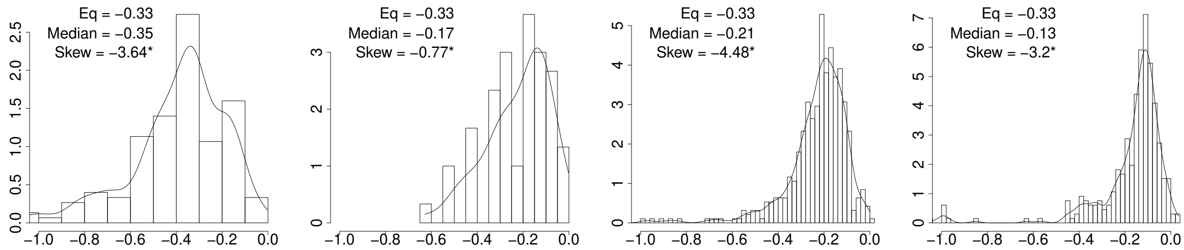
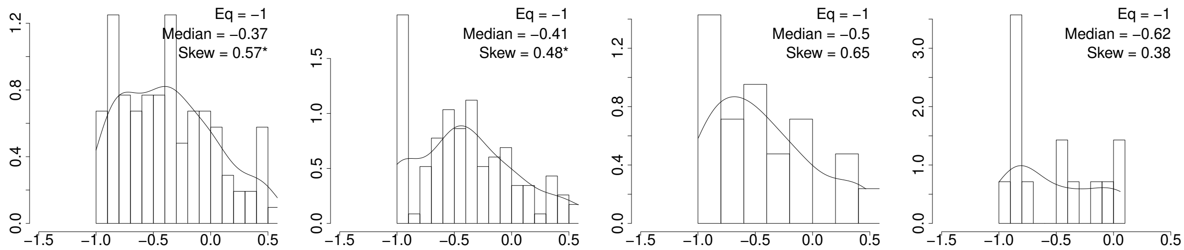


Figure 2: First-price auctions with common values (KL86), inexperienced subjects (a–d) vs. experienced subjects(e–h). Plots are histograms of $r = (Bid - Signal)/w$

(a) $N = 4, w = 6$ (b) $N = 7, w = 6$ (c) $N = 4, w = 12$ (d) $N = 7, w = 12$



(e) $N = 3-4, w = 12, 18$ (f) $N = 5-7, w = 12, 18$ (g) $N = 3-4, w = 24, 30$ (h) $N = 5-7, w = 24, 30$

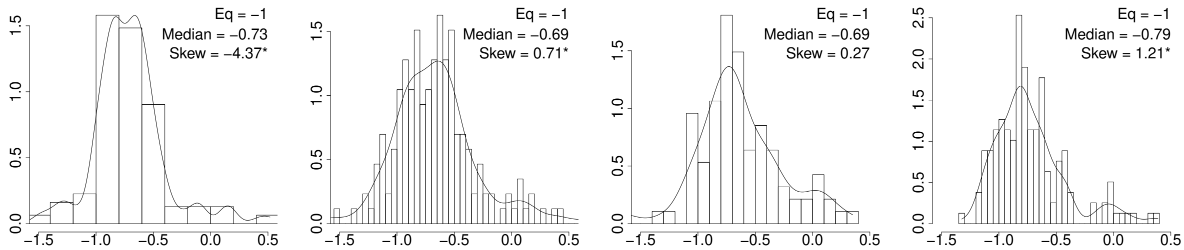


Figure 3: First-price auctions with independent private values (DKL89). Inexperienced subjects (a–c) and experienced subjects (d–f). Histograms of $Bid/Signal$

(a) Inexp: DKL89, $N = 3$ (b) Inexp: DKL89, $N = 6$ (c) Exp: DKL89, $N = 3$ (d) Exp: DKL89, $N = 6$

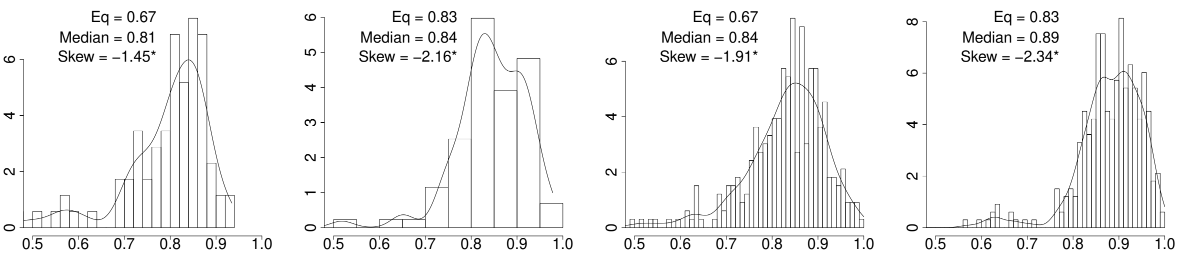


Table 3: Statistical tests of differences in the degree of overbidding and within-subject variance between auctions with affiliated private values and common values

Data	Degree of Overbidding APV		Degree of Overbidding CV	Within-Subject Variance APV		Within-Subject Variance CV	Between-Subj Variance
Inexperienced, $w = 6$	-0.128	$\ll$	0.641	0.359	$\approx$	0.341	0.246
Inexperienced, $w = 12$	0.104	$\ll$	0.523	0.052	$\ll$	0.4	0.147
Inexperienced, all w	-0.058	$\ll$	0.621	0.326	$\approx$	0.361	0.21
Experienced, $w \leq 18$	0.062	$\ll$	0.403	0.159	$<$	0.336	0.248
Experienced, $w \geq 24$	0.179	$<$	0.329	0.113	$\ll$	0.32	0.13
Experienced, all w	0.142	$\ll$	0.357	0.15	$\ll$	0.343	0.162

Description: The sole difference to Table 4 is that the comparison is between APV and CV auctions, instead of inexperienced and experienced subjects.

finite mixture models with up to three components. Each component is characterized by a mean normalized strategy, a between-subject variance regarding the subjects making up the component, and a within-subject variance to capture individual randomization.[12] The impression given by the histograms is confirmed and summarized as follows (the detailed results are not needed further and relegated to the supplementary material).

Result 2. *Across information conditions and experience levels, subject pools consist of single components. Secondary components are either insignificant (16 of 18 treatments) or contain less than 10 percent of the subjects (2 of 18 treatments).*

As a corollary, evidence for multiple discrete levels of reasoning is not obtained.

4.3 First moments of bids: Overbidding

In the standard first-price auctions (see Table 1), normalized BNE bids are $r = -1$ in CV auctions, $r = -2/n$ in APV auctions, and $r = (n-1)/n$ in IPV auctions. The difference between normalized observed bid and normalized BNE bid is called "degree of overbidding". I estimate it controlling for subject-level random effects and evaluate significance using bootstrapping (as before). Table 2, column "Degree of Overbidding", reports the results. The degree of overbidding is significantly positive in 15 out of 18 cases (at $\alpha = .005$), which confirms the usual finding that subjects overbid.

Result 3. *In standard first-price auctions, subjects overbid significantly.*

Comparative statics of overbidding across information conditions can be studied comparing KL86's CV auctions and in KHL87's APV auctions. They implement common values and affiliated private values in otherwise equivalent conditions: signal bandwidths w are similar, numbers of players N are similar, and even experimental instructions

[12] In order to focus on whether discrete components need to be distinguished, the within-subject variances are held fixed constant across components. The models are estimated using the EM algorithm using 25 different starting values in each case, and the number of model components is estimated by maximizing the integrated classification likelihood (ICL), following Biernacki et al. (2000). Maximizing the ICL estimates the correct number of components of finite mixtures more consistently than say Bayes Information Criterion (BIC). See McLachlan and Peel (2000) for further information.

Table 4: Statistical tests of the degree of overbidding and within-subject variance (with respect to the degree of overbidding) as a function of experience

Data	Degree of Overbidding			Within-Subject Variance			Between-Subj
	Inexperienced		Experienced	Inexperienced		Experienced	Variance
Independent private values auctions (DKL89, KL93)							
$N = 3$	0.104	$\approx$	0.148	0.16	$\approx$	0.123	0.033
$N = 5$	0.08	$>$	-0.144	0.142	$<$	0.36	0.077
$N = 6$	-0.021	$<$	0.036	0.164	$\approx$	0.11	0.039
all N	0.05	$\approx$	0.04	0.156	$\approx$	0.178	0.119
all, contr. for N	0.05	$\approx$	0.041	0.155	$\approx$	0.179	0.103
Affiliated private values auctions (KHL87)							
$w = 12$	0.104	$\approx$	0.062	0.051	$\approx$	0.172	0.17
All data	-0.058	$\lll$	0.142	0.331	$>$	0.15	0.134
All, contr. for w	0.058	$\approx$	0.04	0.192	$\approx$	0.209	0.117
Common value auctions (KL86)							
$N \leq 4$, $w \in \{12, 18\}$	0.538	$>$	0.228	0.415	$\approx$	0.343	0.208
$N \leq 4$, all w	0.63	$\ggg$	0.316	0.37	$\approx$	0.363	0.233
$N \geq 5$, all w	0.613	$\ggg$	0.389	0.344	$\approx$	0.309	0.254
all N, $w \in \{12, 18\}$	0.517	$\approx$	0.404	0.397	$\approx$	0.332	0.263
all N, all w	0.621	$\ggg$	0.359	0.357	$\approx$	0.332	0.242
all N, all w, contr. for w	0.573	$\approx$	0.411	0.349	$\approx$	0.337	0.247

Description: The table reports the results of one set of statistical tests per row. Given the subset of data specified in column 1, two null hypotheses are simultaneously tested: (i) H_0 : the degree of overbidding does not differ between inexperienced and experienced subjects, and (ii) H_0 : the residual (i.e. within-subject) variances do not differ between them. These nulls are tested in regression models with the degree of overbidding as independent variable and the level of experience as independent variable (without intercept). $\ggg, \lll$ indicate rejection of H_0 at the .005 level and $>, <$ indicate rejection at .05, where the p-values are bootstrapped as described above. Considering the Bonferroni correction for the multiple testing problem inherent in this analysis, results should be significant roughly at the .005 level. Terms such as the degree of overbidding are used as defined above (e.g. Table 2).

and logistics are similar. The econometric approach is the same as before, regressing the degree of overbidding on the information condition (APV or CV), controlling for subject-level random effects and bootstrapping p-values. Table 3 presents the results: Across conditions and experience levels, the degree of overbidding is significantly higher in CV auctions (5 of 6 times at $\alpha = .005$).

Result 4. *The degree of overbidding is higher in CV auctions than in APV auctions.*

4.4 Second moments of bids: Strategic randomization

As shown above, subjects make either relative or absolute reductions to their signal. They randomize, however, as the positive within-subject variances in Table 2 show. Randomization may be strategic in the sense of mixed equilibrium or erratic in the sense of stochastic choice. Arguably, randomization may be (partially) strategic even if subjects are not aware of their randomization, e.g. "rock-paper-scissors" tends to be perceived as tactical game, but it can be fully erratic only if the within-subject variance decreases as subjects gain experience. Based on this idea, I test for strategic randomization.

Estimates of the within-subject standard deviations are obtained in the regressions above (Table 2), and their interaction with experience is evaluated in models with different within-subject variances for the two levels of experience. Table 4 presents the results in the columns on the "Within-Subject Variance". I test the null hypothesis that variance is constant in multiple ways, either holding the conditions such as number of players N or signal bandwidth w constant, or pooling the data and then controlling for N or w. Overall, the results strongly indicate that the within-subject variance does not change as subjects gain experience. This holds both in treatment-wise comparisons when they are possible, noting that treatment parameters in some experiments are changed as subjects gain experience (see Table 2), and after pooling treatments. Between the 13 tests in Table 4, there is exactly one significant relation for either direction at the .05 level, and none at the .005 level suggested by the Bonferroni correction.

Result 5. *The within-subject variance is constant, suggesting strategic randomization.*

4.5 Third moments of bids: Skewness

Regarding the third moments of the bid distribution, the histograms in Figures 1–3 show that the overall distributions are significantly left-skewed in private value auctions (in both IPV and APV auctions), while skewness tends to be inverted in CV auctions.[13] In all five cases where exact treatment-wise comparisons between inexperienced and experienced subjects are possible, the estimated skewness further shifts toward left-skewed distributions as subjects gain experience. Due to subject heterogeneity, however, the overall skewness may not equate with the average individual skewness. The average individual skewness is the skewness of the errors when regressing the normalized bids on the intercept controlling for subject-level random effects. These estimates, reported in Table 2 in column "Skewness", are similar to the overall skewness estimates: Skewness is mostly significant, at least at $p = .05$, and if it is significant, then toward left-skewness in PV auctions and toward right-skewness in CV auctions.

Result 6. *Distributions of bids are left-skewed in private value auctions and right-skewed in common-value auctions.*

5 Behavioral models in relation to the stylized facts

Table 5 summarizes the relation of the predictions of the models reviewed in Section 2 to the stylized facts just distilled. On this qualitative basis, type projection explains all observations that existing concepts explain, and in addition it explains observations that existing concepts do not explain. Specifically, type projections explains first moments (overbidding) as well as the existing concepts in combination, and it uniquely explains most observations on the higher moments (randomization and skewness).

To illustrate, I discuss the predictions of type projection equilibrium in some detail (for full details, see Breitmoser, 2015).[14] For simplicity, I consider first-price auctions assuming continuity of signals and bids. A player gets signal x, his expectation of the

[13] All of the histograms additionally present information on the skewness of the distributions of normalized bids. An asterisk is printed next to the skewness estimate if it deviates significantly from zero. As above, significance of the skewness is evaluated by bootstrapping, resampling at the subject level.

[14] The predictions of the existing concepts are well-known and therefore skipped here. BNE for expected payoffs by definition does not predict overbidding, risk aversion (CRRA) predicts overbidding in private value auctions, and cursed equilibrium predicts overbidding in common value auctions. All these concepts

Table 5: Stylized facts in relation to the models' predictions

	Empirical Observation	Theoretical prediction				
		Exp payoff	Risk aversion	Cursed Eq	Level-k	Type projection
Distribution	Unimodal	×	×	×		×
Overbidding PV	Yes		×			×
Overbidding CV	Yes			×	(×)	×
Overbidding Comp	CV>PV			×	(×)	×
Individual variance	Const					×
Skewness PV	Left					×
Skewness CV	Right					

Note: The predictions for expected payoffs (Eq. 1), risk aversion (Eq. 5), cursedness (Eq. 4), and type projection (Eq. 3) are for BNE. The predictions for level-k are derived for the standard assumption (Stahl and Wilson, 1995; Nagel, 1995) that level-0 players randomize uniformly given their actual knowledge. Crawford and Iriberri (2007) discuss a level-k model where level-0 players randomize uniformly given "my" knowledge. This model predicts overbidding in CV auctions, which I indicate using parentheses.

object value conditional on x is $v(x)$. The expectation conditional on own signal x and highest opponent signal y is $v(x,y)$. The density of the highest opponent signal y conditional on the own signal x is $f_Y(y|x)$. A pure strategy $b_\star$ is a continuous, monotonic function mapping signals x to bids $b \in \mathbb{R}$. Now, applying Equation (3), the expected payoff of a type-projecting player bidding $b \in \mathbb{R}$, conditional on own signal x and in response to the opponents' bidding function $b_\star$, is

$$\Pi_\rho(b|b_\star,x) = (1-\rho)\int_{\underline{x}}^{b_\star^{-1}(b)} \big(v(x,y)-b\big) f_Y(y|x)\,dy + s\cdot\rho\int_{\underline{x}}^{\overline{x}} \big(v(x,y)-b\big) f_Y(y|x)\,dy. \quad (6)$$

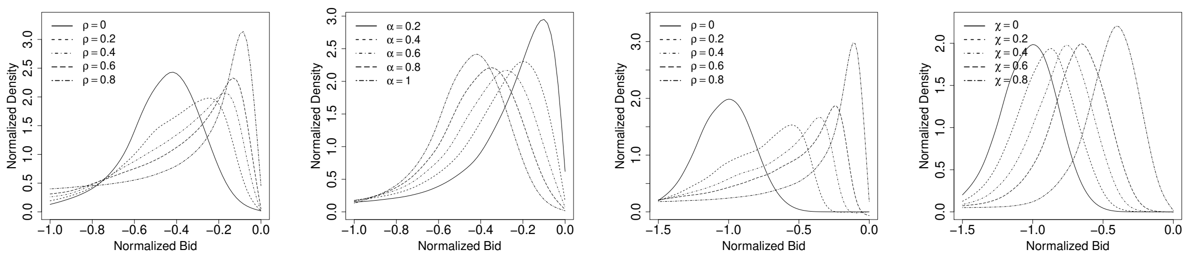
The second summand captures the case of projection. In this case, the expected payoff depends on the relation of b and $b_\star(x)$, with $s=0$ if $b<b_\star(x)$, $s=1/n$ if $b=b_\star(x)$, and $s=1$ if $b>b_\star(x)$. One may think of s as the probability of winning in the case of projection. The optimal bid b maximizes the expected payoff of i, and in any pure equilibrium, $b=b_\star(x)$ obtains. The implications of projection relate to two well-known phenomena, loser regret and broken Bayesian updating.

Loser regret If a projecting player bids less than opponents with the same signal, $b<b_\star(x)$, he underestimates the probability of winning, as he believes to lose with certainty in the projection case. If he bids more than opponents with the same signal, he overestimates the probability of winning, as he believes to win with certainty in the projection case. This induces an incentive to outbid opponents with the same signal, in all information conditions (in first-price auctions). These incentives resemble *loser regret* (Engelbrecht-Wiggans, 1989), i.e. to anticipate regret if a higher bid would have won the auction profitably. Projecting players act as if they anticipate "conditional loser regret", i.e. regret if a higher bid would have won the auction against opponents with the same valuation. The differences are minor, as loser regret materializes only if the opponents' values are similar. Thus, I will say that projection induces loser regret as observed by

predict pure equilibria, which explains neither randomization nor skewness. For detailed discussion on the existing concepts and on level-k, let me refer to Crawford and Iriberri (2007).

Figure 4: Projection predicts skewed overbidding in both APV and CV auctions. Risk aversion and cursedness predict symmetric overbidding in APV and CV, respectively

(a) APV: Projection ρ (b) APV: Risk aversion α (c) CV: Projection ρ (d) CV: Curse χ



Filiz-Ozbay and Ozbay (2007) and Engelbrecht-Wiggans and Katok (2007).[15]

Broken Bayesian updating If one outbids opponents with the same signal, $b > b_\star(x)$, the expected object value under projection is equal to the expectation under cursedness (exchanging ρ and χ). Then, projection attenuates Bayesian updating. If $b < b_\star(x)$, the expected payoff conditional on winning equates with the Bayesian expectation. Hence, the conditional expectation is biased only if one *outbids* opponents with the same value. In standard common value auctions, the bias is an upward bias, i.e. the conditional expectation exhibits an upward jump at $b = b_\star(x)$. Besides inducing cursed expectations, the upward jump adds to the loser regret. Thus, the incentives of projecting players to outbid opponents of the same type are strongest in common value auctions. On a qualitative basis, type projection therefore predicts that if we hold the degree of projection constant, overbidding occurs in both information conditions, but the normalized degree of overbidding is larger in common value auctions than in private value auctions.

Randomization For purpose of contradiction, assume there exists a pure ρ-TPE, $\rho \in (0,1)$, with equilibrium strategy $b_\star$. Any $\varepsilon > 0$ implies that outbidding the opponents by $b = b_\star(x) + \varepsilon$ increases the probability of winning in case of projection (weight $\rho > 0$) discretely from $1/n$ to 1, while it decreases the expected payoff conditional on winning only infinitesimally.[16] If the conditional payoff after bid increment is positive, the projecting player prefers outbidding the opponents to matching their bids. In turn, a symmetric, pure strategy profile can be an equilibrium only if it induces non-positive expected payoffs. Then, however, projecting players can realize positive profits by deviating to bids $b < b_\star(x)$ if $\rho < 1$. Hence, a ρ-TPE must be mixed if bids are continuous.[17]

[15] Note that the projected probability of winning is discontinuous in b if the opponents play a pure strategy. It jumps at $b = b_\star(x)$ where one "overtakes" opponents with the same signal. The discontinuity will disappear once we allow for mixed strategies, but the incentive to slightly outbid opponents with similar values is robust to allowing for mixed strategies.

[16] Without projection, $\rho = 0$, both effects are infinitesimal and thus balanced in a pure BNE.

[17] This prediction applies in all auctions exhibiting strategic complements, i.e. in all auctions considered here (see Breitmoser, 2015), but for example not in second-price IPV auctions.

Skewness Given that it is mixed, it is easy to see that the type projection equilibrium strategy must be left-skewed, i.e. that the density is increasing (under strategic complementarity). Along the support of the equilibrium, a player is indifferent. A projecting player can be indifferent toward increasing his bid only if the increase would reduce his expected payoffs without projection, but the reduction is offset by the increase in the share of opponents perceived to be overtaken due to projection. The latter is the density of the opponents' strategy, which therefore must be increasing. To formalize the argument, consider a two-player APV or CV auction with strategic complements, unconditional value being equal to one's signal x, and an opponent playing a mixed strategy σ randomizing over normalized bids $r < 0$. Let F_σ denote the c.d.f. of σ and let $\tilde{\Pi}(r|\sigma)$ denote one's expected payoff (without projection) of normalized bid r in response to σ. The expected payoff of a projecting player is $\tilde{\Pi}_\rho(r|\sigma) = (1-\rho)\,\tilde{\Pi}(r|\sigma) - r\rho\,F_\sigma(r)$. Taking the derivative with respect to $r < 0$, in the interior of the support, using the fact that it is zero there, and rearranging terms, we obtain

$$\sigma(r) = -\frac{F_\sigma(r)}{r} + \frac{1-\rho}{\rho} \cdot \frac{\tilde{\Pi}'(r|\sigma)}{r}. \tag{7}$$

Now, $\sigma(r)$ is increasing in r, as $F_\sigma(r)$ is increasing and $\tilde{\Pi}'(r|\sigma)$ is decreasing (for details, see Breitmoser, 2015). Figure 4 plots the predictions of type projection in APV and CV auctions, alongside those of risk aversion in APV auctions and cursedness in CV auctions.[18] The predictions are plotted for logit equilibria as analyzed below, which illustrates that the predicted shape of the equilibrium strategies is robust to (small) logit errors. In contrast to projection, risk aversion and cursed equilibrium predict (largely) symmetric distributions even with logistic errors.

6 Toward a robust analysis of bidding

Table 5 suggests that type projection comprehensively explains bidding. This suggestion is tested rigorously next. A test is necessary, as the stylized facts cover complexity and moments of bidding fairly systematically, but their analysis cannot be complete. An analysis of the full distribution, exploiting the properties of maximum likelihood estimators, can complete the picture and show which models indeed capture behavior. Specifically, a model's ability to explain signs of moments ("stylized facts") does not imply that the moments are of the correct scale, that they are captured simultaneously or that they are predicted robustly across conditions (the comparative statics). Joint explanation of moments and comparative statics are tested in structural analyses if we pool treatments and experiments—and in this sense, the structural analysis will put the discussion of model adequacy on a solid econometric basis. Several issues are to be considered, though.

[18] Risk aversion does not affect equilibrium predictions in common value auctions and cursedness does not affect predictions in private value auctions. Hence, the corresponding plots are skipped.

6.1 Stochastic choice

The objective is to identify payoff structure (Eqs. 1–5) and belief systems. All concepts but type projection equilibrium predict pure strategies and thus fit the observations only if we allow for stochastic choice (i.e. "errors"). To not rule out these models right away, and in line with virtually all previous analyses, I allow for stochastic choice. There is no consensus on modeling stochastic choice, however. Analyses in experimental game theory typically allow for "logistic" perturbations of utilities. Given strategic beliefs $\tilde{\sigma}_i$ and payoff structure $\tilde{\pi}_i$, subjects do not choose best responses (Eq. 2) but logit responses

$$Logit_{t_i}(\tilde{\sigma}_{-i}|\tilde{\pi}_i,\lambda) = \big\{\sigma^l(a_i)\big\}_{a_i\in A_i} \quad \text{with} \quad \sigma^l(a_i) = \frac{\exp\{\lambda\,\tilde{\pi}_i\big(a_i|t_i,\tilde{\sigma}_{-i}\big)\}}{\sum_{a_i'\in A_i}\exp\{\lambda\,\tilde{\pi}_i\big(a_i'|t_i,\tilde{\sigma}_{-i}\big)\}}. \tag{8}$$

Logit implies that the higher the expected payoff of an action, the higher its probability, with the precision parameter λ ranging from $\lambda = 0$ (uniform randomization) to infinity (best response, Eq. 2). In contrast, empirical analyses of auctions typically allow for "behavioral" perturbations of choices. Using the notation in Eq. (2),

$$Behav_{t_i}(\tilde{\sigma}_{-i}|\tilde{\pi}_i,\lambda) = \big\{\sigma^b(a_i)\big\}_{a_i\in A_i} \quad \text{with} \quad \sigma^b(a_i) = f\big(a_i - BR_{t_i}(\sigma_{-i}|\pi_i)\big) \tag{9}$$

where $f(\cdot)$ is the density function of the error distribution. Analyses typically estimate the parameters hidden in BR_{t_i} by least squares. Behavioral errors imply that the closer an action to the best response (in the strategy space), the higher its probability. These models have been tested repeatedly, and it appears that choice probabilities relate more closely to utility differences than to distances in strategy space (e.g. McKelvey and Palfrey, 1995, Weizsäcker, 2003, Breitmoser, 2013). For this reason, I assume logistic errors.[19]

6.2 Identifying payoff structure and belief system

The belief systems usually considered in analyses of auctions may be labeled "equilibrium beliefs" (rational expectations), "level-k beliefs" (as defined above), and "Nash beliefs". By Nash belief, I refer to the belief that opponents play the BNE strategies for the respective payoff structure, e.g. cursed equilibrium in the case of cursed payoffs. Note the subtle difference between "equilibrium beliefs" and "Nash beliefs": players with so-called equilibrium beliefs have rational expectations and anticipate errors of opponents (e.g. logistic errors), while players with Nash beliefs do not anticipate errors.

In principle, it is possible to combine any of these belief models with any payoff structure $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$. Indeed, previous analyses examined a fairly large variety of combinations, but unfortunately with little overlap between studies. Goeree et al. (2002b) examine equilibrium beliefs in conjunction with risk aversion (and logit errors). Eyster and Rabin (2005) examine cursed equilibrium, i.e. Nash beliefs in conjunction with cursed payoffs (and behavioral errors). Crawford and Iriberri (2007)

[19] In turn, models of behavioral errors are rather tractable if one additionally assumes that subjects do not anticipate the errors of their opponents (Bajari and Hortacsu, 2005, discuss some differences in tractability). Subjects not anticipating errors consequently anticipate the BNE strategies. In the presence of errors, this violates rational expectations, but it simplifies estimation as only the BNE strategies need to be computed, not the equilibrium strategies accounting for errors.

show that level-k beliefs fit better than Nash beliefs in private and common value auctions, assuming either expected or cursed payoffs and logit errors. Bajari and Hortacsu (2005) show that Nash beliefs with behavioral errors fit about as well as equilibrium beliefs with logit errors (in the sense that the differences are insignificant). That is, the only model considered by two studies is equilibrium beliefs with logit errors, considered in Goeree et al. (2002b) and Bajari and Hortacsu (2005), and in turn, the existing results do not form a comprehensive picture of the behavioral forces underlying bidding.

In addition, the studies show that the identified payoff structure depends on the assumed belief system and vice versa. This relates to analyses of choice under risk, where identification of utility functions and probability weighting depends on the model of stochastic choice, see e.g. Hey (2005), Blavatskyy and Pogrebna (2010), and Wilcox (2011). Thus, to reliably analyze the payoff structure, we have to relax the assumptions on belief formation and simultaneously allow for all three of the standard models—loosely speaking, to let the data decide. To illustrate how this is achieved, let me define quantal response equilibrium (McKelvey and Palfrey, 1995) with logit errors.

Definition 1 (Quantal response equilibrium)**.** Given $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a λ-QRE if all types $t_i \in T_i$ of all players $i \in N$ choose $\sigma_i(\cdot|t_i) = Logit_{t_i}(\sigma_{-i}|\tilde{\pi}_i, \lambda)$.

QRE is the standard model in behavioral game theory, and successfully captures behavior in e.g. the centipede game (Fey et al., 1996), the traveler's dilemma (Capra et al., 1999), public goods games (Goeree et al., 2002a), monotone contribution games (Choi et al., 2008), and beauty contests (Breitmoser, 2012). Now assume that players believe their opponents play QRE, i.e. their belief is strategically balanced but noisy, and to their belief, they logit respond with their own, presumably higher precision.

Definition 2 (Asymmetric quantal response equilibrium)**.** Given a payoff structure $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is a (λ, κ)-AQRE if there exists a κ-QRE σ' such that for all $t_i \in T_i$ of all $i \in N$, $\sigma_i(\cdot|t_i) = Logit_{t_i}(\sigma'_{-i}|\tilde{\pi}_i, \lambda)$.

By AQRE, players λ-logit respond to a κ-QRE.[20] AQRE nests the three models discussed above, QRE for $\kappa = \lambda$, Level-1 for $\kappa = 0$, and logit response to Nash beliefs for $\kappa = \infty$, and additionally allows for a continuum in-between these extremes. Thus, AQRE is indeed flexible enough to let the data speak for itself, and in addition, it is parsimonious, nesting the three models by adding just one parameter.

The main difficulty with using AQRE is that the underlying QRE needs to be computed explicitly.[21] This is not generally straightforward, as mixed bidding functions are rather complex, but thanks to the above result that subjects' strategies are one-dimensional, AQRE is computationally feasible using current technology.[22] Thus, the above result on complexity allows us to endogenize the belief assumptions made in the literature.

[20]AQRE differs from the asymmetric logit equilibrium defined by Weizsäcker (2003) only insofar as opponents do not know that I use some $\kappa \neq \lambda$. Here, they simply play the QRE with precision λ.

[21]As indicated already, the insight of Bajari and Hortacsu (2005) allowing to avoid the fixed point computation underlying QRE—by exploiting rational expectations and using observed behavior as beliefs—is infeasible here. By assumption, observed behavior forms an AQRE and subjects do not have rational expectations. Thus, the fixed point defining QRE needs to be computed explicitly.

[22]For illustration, consider again a grid with 100 different normalized bids over which the bidders randomize. In an auction with 5 bidders and say 100 possible signals, evaluating the payoff function is possi-

As discussed next, addressing the concerns voiced by Haile et al. (2008) that a sufficiently generalized QRE can fit everything, I verify the fallacy to overfitting by examining predictive adequacy. As robustness checks, I additionally consider the best known alternative belief models, namely level-k (Stahl and Wilson, 1995; Nagel, 1995), cognitive hierarchy (Camerer et al., 2004) and noisy introspection (Goeree and Holt, 2004). These models are defined in Appendix A and discussed briefly below.

6.3 Robustly measuring model adequacy

I report results for three measures of model adequacy. The *descriptive adequacy* measures goodness-of-fit in-sample, the *predictive adequacy* measures the reliability of predictions across experiments, and the *inferential adequacy* measures accuracy of the object values (signals) inferred from bids. Generally, the measure to be used depends on the purpose of the analysis, and thus reporting all three measures may help clarify which model is suitable for which purpose. The main purpose here is to identify payoff structure and beliefs in a manner ensuring that they can be used reliably in future work. To this end, predictive adequacy is most appropriate.

To clarify, let me introduce some notation. D_e denotes the data set associated with experiment e, $D = \cup_e D_e$ denotes the pooled data, and define $D_{-e} = D \setminus D_e$ (the data sets used here are listed in Table 1). Given a model, p denotes a generic parameter vector and $p^\star(D')$ denotes the maximum likelihood estimate given data set D'. Finally, $|p|$ denotes the the dimensionality of p, $|D'|$ denotes the number of subjects in data set D', and $ll(p|D')$ denotes the log-likelihood of the model with parameters p given data D'.

As usual, I measure descriptive adequacy by Bayes information criterion (Schwarz, 1978), using the number of subjects as number of observations.

Definition 3 (Descriptive adequacy). $BIC = -ll\big(p^\star(D)\,|\,D\big) + |p^\star(D)|/2 \cdot \log|D|$

Descriptive adequacy is informative, as a high compatibility with stylized facts does not guarantee descriptive adequacy. The pooled data, including scale of moments and comparative statics, need to be explained using a single parameter vector.

Second, I measure predictive adequacy by fitting the parameters to one information condition, using the estimate to predict the remaining data, and rotating such that all data sets are used as training data.[23] There is no penalty term as in BIC, as by definition no parameter is adjusted to the data set used in the validation stage, i.e. no degree of freedom is used. To be aligned with the other measures, I report the absolute values of the log-likelihoods, which implies that "less is better" for all measures.

ble by simulation using quasi-random numbers, which in turn can be implemented in a massively parallel manner on a GPU (which is reasonably standard, see e.g. Breitmoser, 2012). On top of it, finding the fixed point for the distribution over 100 normalized bids is possible, but finding the fixed point for 100×100 probabilities is not (yet) possible using "regular" computers.

[23]The tendency to distinguish descriptive and predictive adequacy is a rather recent development in analyses of decision-theoretic models (Wilcox, 2008; Hey et al., 2010), learning models (Erev and Roth, 1998; Camerer and Ho, 1999; Tang, 2003; Ho et al., 2008), and simple games (Blanco et al., 2011; Shapiro et al., 2014). I am not aware of existing analyses in Bayesian games in general or auctions in particular. The approach adopted here may be called cross-validation (Browne, 2000) with nonrandom holdout samples (Keane and Wolpin, 2007).

Definition 4 (Predictive adequacy). $LL_{pred} = -\sum_e ll\big(p^\star(D_e)\,|\,D_{-e}\big)/(n-1)$ using n as number of experiments analyzed

Predictive adequacy assesses the reliability of estimates in other experiments, and this way, it assesses robustness of the behavioral relevance of a motive. In general, high descriptive adequacy does not imply predictive adequacy (see also Hey et al., 2010), as predictive adequacy requires goodness-of-fit not just in the ex-post sense, after fitting the model to the data, but also in an ex-ante sense, after fitting the model to different data. This allows me to address Haile et al. (2008), who show that in any game with $K \geq 2$ actions, a generalized logit model with $K-1$ parameters can fit any choice pattern. Given this dimensionality, any model should be able to fit perfectly, but in analyses of auctions, hundreds of parameters would be required and parameters would have to be treatment-specific. Nonetheless, by verifying predictive adequacy, I can explicitly address the concerns of Haile et al. and transparently evaluate (lack of) reliability.

Finally, the inferential adequacy also is evaluated out-of-sample, but now, we infer signals from bids rather than predicting bids from signals (following Bajari and Hortacsu, 2005). Given an observation and a set of parameters (estimated using training data), the theoretical bidding function for the respective out-of-sample treatment is determined and the expectation of the signal conditional on the observed bid is computed. This conditional expectation is called *inferred signal*. The inferential adequacy is the mean absolute deviation (MAD) to the actual signals. The supplementary material additionally lists the results for the mean squared deviation (MSD), which are very similar. Formally, let $m(p|D')$ denote the mean absolute deviation of inferred signals to actual signals if inference is made using parameter vector p on data set D'.

Definition 5 (Inferential adequacy). $MAD = \sum_e m\big(p^\star(D_e)\,|\,D_{-e}\big)$

Inferential adequacy is desirable in empirical work (Bajari and Hortacsu, 2005) and complements predictive adequacy by its focus on the *expectation* of the underlying signal. Specifically, the first moment of its distribution needs to be predicted, rather than the full distribution as in likelihood-based procedures.

7 Analysis of bidding in auctions

The analysis distinguishes inexperienced subjects (first five auctions, following Crawford and Iriberri, 2007) and experienced subjects (last five auctions), and throughout I report the results for all three measures of model adequacy. The objective to obtain results useful in future work suggests to use predictive adequacy, which also evaluates potential overfitting, but the results are robust to the chosen measure of model adequacy. The levels of significance used above, 0.05 and 0.005, are used equally here, and the likelihood ratio tests again are bootstrapped (using nested or non-nested Vuong tests, as appropriate).

Initially, I analyze the "standard" auctions in Table 1, i.e. auctions with common value (CV, Kagel and Levin, 1986), affiliated private values (APV, Kagel et al., 1987), or independent private values (IPV, Kagel and Levin, 1993; Dyer et al., 1989). Given the magnitude of the between-subject variances in Table 2, I allow for heterogeneous subjects as described in Appendix A. The appendix also specifies the (standard) likelihood

Table 6: Analysis of the payoff structure (for all measures of adequacy: less is better)

(a) Inexperienced subjects (first five auctions), across all standard auctions (IPV, APV, CV)

Adequacy	Exp Payoff		Risk Aversion		Projection		Cursedness
Descriptive	4108	$\gg$	3911	$\gg$	3742	$\ll$	3967
Predictive	4261	$\approx$	4203	$\gg$	4034	$\ll$	4449
Inferential	2226	$\ll$	2636	$\gg$	1759	$\ll$	2110
Average (λ,κ,α)	$45, 0.12$		$0.39, 15, 0.34$		$11, 3.1, 0.44$		$20, 1.3, 0.73$

(b) Experienced subjects (last five auctions), across all standard auctions (IPV, APV, CV)

Adequacy	Exp Payoff		Risk Aversion		Projection		Cursedness
Descriptive	4005	$\gg$	3573	$\gg$	3377	$\ll$	3805
Predictive	4069	$\gg$	3799	$>$	3686	$\ll$	4200
Inferential	4460	$>$	4004	$\approx$	3498	$\approx$	3650
Average (λ,κ,α)	$47, 0.05$		$0.34, 14, 0.24$		$18, 3.3, 0.48$		$130, 0.21, 0.78$

Note: The payoff structures are defined in Eqs. (1)–(5) and the measures of model adequacy are defined in Definitions 3–5. The row "Average (λ,κ,α)" lists the average estimates of precision λ, belief parameter κ (of AQRE), and degree α of risk aversion/cursedness/projection (depending on model). Significance at 0.05 is indicated by $<,>$, and significance at 0.005 is indicated by $\ll,\gg$ (which implements the Bonferroni correction for nine simultaneous tests per level of experience).

function and the numerical approach to its maximization.[24] Robustness checks of belief systems and homogeneous subject pools are discussed below, while the full set of details and all parameter estimates are provided as supplementary material.

7.1 Which payoff structure is most adequate?

The analysis proceeds sequentially. First, I analyze the payoff structure, defined in Eqs. (1)–(5), under the general model of strategic beliefs (AQRE, Def. 2) nesting the three standard models: rational expectations, naive beliefs, and Nash beliefs.

Question 1. *Allowing for all of the standard belief systems, which payoff structure captures bidding: expected payoffs, risk aversion, cursedness, or projection?*

The results are presented in Table 6. The overall picture is rather clear-cut: Type projection generally is most adequate, corroborating the compatibility with the stylized facts, and in most cases the differences to the other payoff structures are highly significant. The descriptive adequacy of projection shows that a fairly constant degree of projection fits behavior across conditions. Its predictive adequacy shows that even if fitted to single data sets, the estimated degree of projection is comparably robust and enables meaningful predictions of behavior in other conditions. This indicates that type projection itself has a robust impact on behavior and that its descriptive adequacy is not due to overfitting.

The predictive adequacy also uncovers the most striking differences to the other concepts. Both risk aversion and cursedness significantly improve on expected payoffs

[24] Briefly, all parameters (precision λ, belief asymmetry κ, and degrees of risk aversion, cursedness, and projection) are randomly distributed across subjects, while each subject is assumed to bid according to a constant set of parameters. The latter is consistent with the lack of time trends found above, for both inexperienced subjects (first five auctions) and experienced subjects (last five auctions).

descriptively (i.e. in-sample) but fail to consistently improve on it predictively (out-of-sample). Their behavioral content in relation to expected payoffs is not robust. In turn, type projection does not only yield higher predictive adequacy than expected payoffs for both inexperienced and experienced subjects—it fits better out-of-sample than expected payoffs does in-sample. Thus, type projection is of robust relevance in bidding. The results on inferential adequacy are similar, though not quite significant in all cases.

Result 7. *Type projection is the dominant model of the payoff structure. It is most adequate by all measures, for both experienced and inexperienced subjects, and it uniquely improves on expected payoffs out-of-sample.*

7.2 How are beliefs formed?

The result that type projection is an adequate representation of the payoff structure, both under the generalized belief system nesting the usual beliefs (shown above) and under other belief systems (shown in the supplementary material), can now be used to identify the belief system itself. Besides equilibrium (QRE) and asymmetric QRE (Definitions , Def. 1 and 2), I will consider noisy introspection (NI, Goeree and Holt, 2004), cognitive hierarchy (CHM, Camerer et al., 2004), and level-k (Nagel, 1995; Stahl and Wilson, 1995), see Definitions 6-8 in Appendix A. The procedure is otherwise consistent with the one used to answer Question 1.

Question 2. *Given the identified model of the payoff structure (type projection), which belief system captures bidding?*

First, to provide context, let me briefly review existing results. In small normal-form games with dominated strategies, subjects exhibit low depth of reasoning: They do not choose dominated strategies but fail to take into account that opponents reason similarly (Costa-Gomes et al., 2001; Weizsäcker, 2003; Costa-Gomes and Weizsäcker, 2008). In games without dominated strategies, in particular in games with unique mixed equilibria, equilibrium beliefs are most adequate (Goeree et al., 2003; Brunner et al., 2011). In large normal form games, beliefs tend to be in-between these extremes: subjects may underestimate the precision of others, but not as extremely as level-1 (Goeree et al., 2002a; Costa-Gomes and Crawford, 2006; Breitmoser, 2012). This can be captured by e.g. AQRE with $\lambda > \kappa > 0$ and NI with $1 > \kappa > 0$. Auctions are similarly large games, and following Goeree et al. (2002b) equilibrium beliefs are adequate (i.e. QRE). Bajari and Hortacsu (2005) show that equilibrium beliefs are about as adequate as Nash beliefs.[25]

The results of the current analysis, provided in Table 7, largely corroborate these observations. To organize the results, let us take the unique one-parametric model (QRE) as benchmark and ask which of the two-parametric models improve on it consistently. As for inexperienced subjects, the only model that improves on QRE consistently (by all three measures) is AQRE, but in two of the three cases, the significance of the differences is not robust to the Bonferroni correction. Thus, I say that AQRE weakly improves on QRE for inexperienced subjects. As for experienced subjects, no model consistently improves on QRE, which has the highest predictive adequacy and thus fits most robustly.

[25]In turn, Crawford and Iriberri (2007) show that for a specific assumption of level-0 behavior and assuming expected payoffs, level-k models may fit better than equilibrium beliefs.

Table 7: Analysis of the belief systems (for all measures: less is better)

(a) Inexperienced subjects (first five auctions), across all standard auctions (IPV, APV, CV)

Adequacy	Level-*k*		CHM		QRE		AQRE		NI
Descriptive	3760	$\approx$	3762	$\approx$	3771	$>$	3742	$\approx$	3751
Predictive	4091	$>$	4014	$\ll$	4082	$>$	4034	$\approx$	4035
Inferential	1962	$\ll$	2210	$\gg$	2110	$\gg$	1759	$\ll$	2116
Average (λ, κ, ρ)	$43, 4.8, 0.44$		$47, 7.1, 0.44$		$45, 0.4$		$11, 3.1, 0.44$		$16, 0.54, 0.5$

(b) Experienced subjects (last five auctions), across all standard auctions (IPV, APV, CV)

Adequacy	Level-*k*		CHM		QRE		AQRE		NI
Descriptive	3404	$\ll$	3435	$\approx$	3406	$>$	3377	$\ll$	3454
Predictive	3644	$\ll$	3697	$\gg$	3599	$\ll$	3686	$\gg$	3609
Inferential	3370	$\ll$	3565	$\approx$	3508	$\approx$	3498	$>$	3251
Average (λ, κ, ρ)	$29, 11, 0.42$		$29, 8, 0.42$		$52, 0.45$		$18, 3.3, 0.48$		$16, 0.62, 0.52$

Note: The belief models are defined in Definitions 1, 2, 6–8, and the measures of model adequacy are defined in Definitions 3–5. The row "Average (λ, κ, ρ)" lists the average estimates of precision λ, belief parameter κ (depending on model), and degree ρ of projection (depending on model). Significance at 0.05 is indicated by $<, >$, and significance at 0.005 is indicated by $\ll, \gg$.

Result 8. *Inexperienced subjects tend to underestimate the precision of others. Beliefs converge to rational expectations (QRE) as subjects gain experience.*

Thus, in line with the literature, inexperienced bidders exhibit comparably noisy beliefs (relating to Crawford and Iriberri, 2007), though level-*k* is not the most adequate model, which confirms the observations that bid distributions are unimodal (Result 2). Experienced bidders are well described holding equilibrium beliefs (relating to Goeree et al., 2002b, and Bajari and Hortacsu, 2005). To illustrate the goodness-of-fit, Figure 5 plots the predicted densities of QRE with projection over the histograms of normalized bids. These plots refer to inexperienced subjects; the respective plots for experienced subjects are similar and provided as supplementary material.

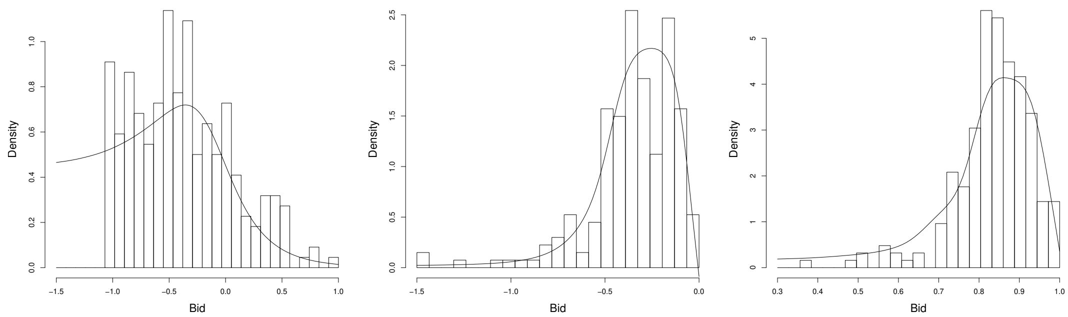
7.3 Further experience effects

QRE fits behavior of experienced subjects most robustly and fits only weakly worse than AQRE for inexperienced subjects. Therefore, I focus on QRE (with projection) to analyze behavioral differences between inexperienced and experienced subjects. Similar analyses for the other belief schemes are provided in the supplementary material.

Parameter estimates As described above, the models evaluated so far allow for unobserved heterogeneity: QRE-precision and degree of projection are constant for each subject but distributed randomly in the population. Table 8 presents their mean values for each of the data sets and for the pooled data set, separately for inexperienced and experienced subjects. In all information conditions, the mean precision increases as subjects gain experience, to the extent that behavior largely converges to projection equilibrium without errors in private value auctions. The degree of projection is on average constant, slightly increasing for private value auctions and substantially decreasing for common

Figure 5: The predictions of QRE with projection (solid lines) in relation to the data

(a) CV, Inexperienced (b) APV, Inexperienced (c) IPV, Inexperienced



Note: The histograms plot the normalized bids for each information condition in standard auctions (see Table 1), always aggregated across treatments, focusing on inexperienced subjects. On top, the solid line depicts the predicted choice probabilities of QRE with projection, equally averaged across treatments.

value auctions. In the latter case, the degree of projection is initially very high ($\rho = 1$) but declines to one of the lowest values across conditions as subjects gain experience. The high initial value indicates that inexperienced subjects struggle comprehending common values, and the subjects struggling the most actually go bankrupt in CV auctions. Bankrupt subjects are removed from the experiment and therefore not present in the pool of experienced subjects, which slightly biases the average degree of projection in CV auctions in relation to the other auctions. Aside from that, the mean degree of projection is near 0.5, which indicates that type projection is indeed a constant factor in bidding.

Table 8: Average precision and degrees of projection in standard auctions

Subjects	CV, 1st	CV, 2nd	APV	IPV	Pooled
Inexperienced	1.7, 1	30, 0.58	18, 0.3	26, 0.42	45, 0.4
Experienced	8.2, 0.28		45, 0.65	41, 0.46	52, 0.45

Note: The average parameters are reported for QRE with projection, listing the QRE precision first and the degree of projection second. "CV, 1st" and "CV, 2nd" indicate first and second price auctions (respectively) with common value, see Table 1.

Are subjects heterogeneous? Next, I investigate how the extent of subject heterogeneity depends on the level of experience. Contrary to the heterogeneous model considered so far, in the homogeneous model, subjects are collectively described by a representative agent with "average" precision λ and "average" degree of projection ρ. The procedure is otherwise equal to above. The results are presented in Table 9. As for inexperienced subjects, allowing for heterogeneity improves the goodness-of-fit descriptively (in-sample), but neither predictively nor inferentially. In this case, allowing for heterogeneity induces overfitting. As for experienced subjects, allowing for heterogeneity highly significantly improves on the representative-agent model according to all three measures. This complements the earlier finding that experienced subjects exhibit higher precision and rational expectations, suggesting that experienced subjects understand auctions, including their opponents, which allows them to actually bid according to their preferences.

Table 9: Analysis of significance of subject heterogeneity

(a) Inexperienced subjects (first five auctions)

Adequacy	Homog.		Heterog.
Descriptive	3893	$\gg$	3771
Predictive	4075	$\approx$	4082
Inferential	2108	$\approx$	2110
Average pars	$17, 0.43$		$45, 0.4$

(b) Experienced subjects (last five auctions)

Adequacy	Homog.		Heterog.
Descriptive	3570	$\gg$	3406
Predictive	4030	$\gg$	3599
Inferential	4613	$\gg$	3508
Average pars	$11, 0.61$		$52, 0.45$

Note: The tables report results for QRE with projection, assuming either homogeneous or heterogeneous subjects (the latter as above and as described in Appendix A). Significance at 0.05 is indicated by $<, >$, and significance at 0.005 is indicated by $\ll, \gg$.

Result 9. *As subjects gain experience, their average precision increases, the average degree of projection remains largely constant, and subjects exhibit heterogeneity.*

7.4 Are multi-motive models more adequate?

Projection and cursedness as defined in Eqs. (3) and (4) can be complemented by any outcome-based utility function. In auctions, is is natural to consider risk aversion (CRRA). On the one hand, examining type projection in relation to a model merging type projection and risk aversion allows us to verify to which degree risk aversion complements projection. On a qualitative basis, type projection explains overbidding in private value auctions as well as risk aversion, and in addition it explains randomization and skewness. It is not obvious that risk aversion covers facets of behavior neglected by projection. On the other hand, examining type projection in relation to a model merging cursedness and risk aversion illustrates to which degree projection (with a given degree of projection) covers the facets explained by risk aversion and cursedness with independent degrees of risk aversion and cursedness, respectively.

The analytical approach is as before. Based on the above results, I focus on heterogeneous subjects and equilibrium beliefs, but comprehensive robustness checks are provided as supplementary material. The results, reported in Table 10, are clear and can be summarized succinctly. The in-sample differences are small and insignificant, i.e. type projection describes behavior comprehensively and does not miss out on any aspect captured by the other models despite its relative parsimony (corroborating the stylized facts compiled in Table 5). The predictive adequacy significantly improves with type projection on its own, indicating that its parsimony indeed improves robustness. Augmenting type projection by risk aversion improves the inferential adequacy when subjects are experienced, which may be of relevance in empirical work.[26]

Result 10. *Multi-motive models do not improve on type projection in-sample (descriptively) or out-of-sample (predictively). Complementing projection by risk aversion improves inferential adequacy for experienced subjects.*

[26]One caveat is that large and diverse data sets are required to reliably estimate both degree of projection and degree of risk aversion (risk aversion on its own lacks inferential adequacy). This seems to be the case in the present analysis but is unlikely to be satisfied in field work. Another caveat is that the models considered here are estimated by maximum likelihood, and thus inferential adequacy is a side effect. If inferential adequacy is the main objective, a different estimator may be appropriate.

Table 10: Evaluation of models with multiple motives

(a) Inexperienced subjects (first five auctions), across standard auctions (IPV, APV, CV)

Adequacy	Proj + RA		Projection		Curse + RA
Descriptive	3772	$\approx$	3771	$\approx$	3790
Predictive	4160	$\gg$	4082	$\ll$	4235
Inferential	2127	$\approx$	2110	$\ll$	2312
Average pars	$46, 0.37, 0.88$		$45, 0.4$		$20, 0.73, 0.64$

(b) Experienced subjects (last five auctions), across standard auctions (IPV, APV, CV)

Adequacy	Proj + RA		Projection		Curse + RA
Descriptive	3378	$\approx$	3406	$\approx$	3424
Predictive	3732	$\gg$	3599	$\ll$	3762
Inferential	2916	$\ll$	3508	$\approx$	3209
Average pars	$67, 0.31, 0.71$		$52, 0.45$		$68, 0.72, 0.46$

Note: The tables report results for QRE with projection. The order of average parameters is (λ, ρ, α) for "Proj + RA", (λ, ρ) for "Projection", and (λ, χ, α) for "Curse + RA", where λ is the QRE-precision, and α, ρ, χ are the degrees of risk aversion, projection, and cursedness, respectively. Significance at 0.05 is indicated by $<, >$, and significance at 0.005 is indicated by $\ll, \gg$.

7.5 Do results also hold in non-standard auctions?

As previewed in Section 2, the auctions labeled "non-standard" are the IPV auction of Goeree et al. (2002b, GHP02) and the CV auction of Avery and Kagel (1997, AK97).

One the one hand, GHP02's auction is labeled non-standard as both signals and bids are discrete, restricted to integers ranging from 0 to 12. This framework may be challenging for type projection, as projecting bidders have a strong incentive to *infinitesimally* outbid opponents with similar values. The incentive to outbid is weaker if discrete bid increments are required, but it is existent and thus qualitatively compatible with overbidding as observed by GHP02. This raises the question if the degree of projection is sufficiently similar between discrete and continuous auctions to consider projection a robust factor in bidding. On the other hand, AK97 consider an auction where the individual signals X_1, X_2 are each uniform on $[1,4]$ and the common value is $X_1 + X_2$.[27] This auction is labeled non-standard as signals conditional on value are not independent. It is a second-price auction, and thus the loser regret implied by projection in first-price auctions is switched off. Projection still implies cursed value perception, i.e. it is largely similar to cursed equilibrium, and thus both concepts are compatible with AK97's observation that players with signals below average overbid and players with signals above average underbid in relation to the BNE.[28] This raises the question if the missing loser regret in

[27] Turocy (2008) discusses the value structure in detail. Such common values arise if one might find oneself in a position to sell the object later (thus, the opponent's value matters) or due to prestige effects.

[28] A difference remains: type projection equilibria are mixed. For contradiction, assume a pure equilibrium and consider a player i with a below-average signal $x_i < 2.5$. If i marginally outbids opponents with the same signal, the expected value conditional on winning is the cursed value $V = \rho(x_i + 2.5) + (1-\rho)(2x_i)$, and if he marginally underbids opponents with the same signal, the conditional value equates with the Bayesian expectation without projection $V = 2x_i$ (both as shown above). Thus, one's conditional expectation is not continuous at the opponent's bid, rules out the existence of pure equilibria.

Table 11: Overview of model adequacy in non-standard auctions

(a) Inexperienced subjects (first five auctions)

	Exp Payoff		Risk Aversion		Projection		Cursedness
AK97	553	$\ll$	556	$\gg$	534	$<$	542
AK97 $\rightarrow$ Standard	4269	$>$	4255	$\gg$	4040	$\approx$	4080
Average pars	3.5		3.2, 1		6.9, 1		5.3, 1
GHP02	694	$\gg$	421	$<$	439	$\ll$	700
GHP02 $\rightarrow$ Standard	4252	$\gg$	3979	$\gg$	3842	$\ll$	4203
Average pars	1.2		13, 0.44		25, 0.5		1.3, 0.51

(b) Experienced subjects (last five auctions)

	Exp Payoff		Risk Aversion		Projection		Cursedness
AK97	524	$\approx$	525	$\approx$	515	$\approx$	517
AK97 $\rightarrow$ Standard	4214	$\gg$	4026	$\gg$	3647	$\ll$	3888
Average pars	37		11, 0.99		16, 0.45		14, 0.59
GHP02	631	$\gg$	354	$\approx$	349	$\ll$	627
GHP02 $\rightarrow$ Standard	4261	$\gg$	3826	$\gg$	3474	$\ll$	4128
Average pars	1.6		11, 0.54		97, 0.35		2.1, 0.51

Note: The rows labeled "AK97" and "GHP02" refer to the descriptive adequacy of the models with respect to the data of Avery and Kagel (1997) and Goeree et al. (2002b), respectively. The average parameters estimated for these data sets are provided in the rows labeled "Average pars". The rows labeled "AK97 $\rightarrow$ Standard" and "GHP02 $\rightarrow$ Standard" refer to the predictive adequacy using estimates from the non-standard auctions to standard auctions. Significance at 0.05 is indicated by $<, >$, and significance at 0.005 is indicated by $\ll, \gg$ (which implements the Bonferroni correction).

this second-price auction and the occurrence of underbidding (for high values) implies that the degree of projection differs between AK97's auction and standard auctions to the extent that it cannot be considered a robust factor in bidding.

The procedure is largely similar to above. Besides descriptive adequacy, Table 11 reports the accuracy of predicting behavior in standard auctions using estimates from either non-standard auction. The remaining measures of adequacy yield similar results and are reported in the supplementary material. The differences in-sample (descriptive adequacy) are small, but differences are large out-of-sample: Projection predicts very well, showing that it is of constant relevance across standard and non-standard auctions. Notably, the estimated degrees of projection are similar to those in standard auctions for both inexperienced subjects and experienced subjects, respectively, and using estimates from GHP02's discrete private auctions, type projection predicts behavior in standard auctions better than the other models capture it in-sample.

Result 11. *In non-standard auctions, type projection also fits best, and in addition, it occurs to the same degree in standard and non-standard auctions.*

8 Conclusion

The purpose of the paper was to analyze payoff structures and belief systems in auctions, in a manner ensuring that the results are robust and applicable in future analyses. The initial hypothesis was that previous analyses yield conflicting results because they use different assumptions on strategic beliefs and possibly omit a common behavioral factor. The former was resolved using a simple model of strategic beliefs nesting the models typically used in analyses of auctions. Regarding the latter, type projection was an ex-ante plausible candidate to be a behavioral factor, as it is robustly observed in psychological research and intuitively applies to all Bayesian games with symmetric type sets—such as auctions. Yet, despite the large amounts of studies dedicated to either, auctions in economics and projection in psychology, the only published paper mentioning a potential link between bidding and projection appears to be Engelmann and Strobel (2012).

The results are consistent and clear. Type projection represents a constant factor in bidding, by its relation to the stylized facts and by both descriptive and predictive adequacy to capture behavior. This holds across information conditions and across levels of experience, and the degree of projection also is largely invariant—while subjects' beliefs approach rational expectations and their precision in maximizing utility increases with experience. Finally, projection is comprehensive, as e.g. complementing it by risk aversion does not improve model adequacy. Thus, there are good reasons to consider type projection a factor of behavior in auctions, and by extension in type-symmetric Bayesian games, which suggests ample opportunity for further research.

In this regard, three points may be worth noting. First, projection likely affects behavior not only in auctions, but similarly in other Bayesian games with symmetric type sets, including games where social preferences matter. In general, though, experimental work in economics tends to attribute deviations from Nash equilibrium either to preferences, such as risk aversion or inequity aversion, or to belief asymmetry, such as level-k. Intuitively, each of these influences affects behavior in general, but projection should not be neglected simply because the literature focused on other concepts so far: Judging by the psychological evidence, the relevance of projection appears to be rather universal.

Second, analysts of empirical auctions may consider projection at least alongside risk aversion in econometric analyses of bidding. This has both a downside and an upside. On the downside, projection equilibria are mixed and their computation may require information that analysts do not immediately have, e.g. the upper bound of values in private value auctions. Less information is required, and some tractability is gained (see Bajari and Hortacsu, 2005), if one is willing to neglect projection and assume "Nash beliefs" (players' beliefs are equilibrium strategies without errors). These assumptions are highly debatable, though. My results challenge the neglect of projection, and most analyses, including Crawford and Iriberri (2007) and above, show that subjects tend to underestimate the precision of others, i.e. the opposite of Nash beliefs. Further on the upside, projection equilibria fit much more robustly than received models across private and common values, which suggests that they are less prone to misspecification of the information conditions. This is promising as many empirical auctions take place in hybrid conditions (Haile, 2001; Goeree and Offerman, 2002). These advantages may well outweigh the additional computational burden, but more work clearly is required.

Finally, Engelmann and Strobel (2012) have shown that subjects are less likely to project if they are provided with the objective information in the best possible way. This suggests that the fallacy to projection may be subject to policy intervention, and future work may determine the best way of providing objective information. Further, to the degree that overbidding is due to risk aversion, educating subjects does not help efficiency. To the degree that overbidding is due to projection, educating subjects increases the efficiency in at least two ways: Subjects stop randomizing in equilibrium, which ensures that the bidder with the highest value wins, and in cases where not just the winners pay their bids (e.g. contests), a reduction of overbidding increases efficiency. Thus, the above findings also have novel policy implications.

A Relegated definitions

Belief models As described above, I endow all models with logistic errors. Noisy introspection (Goeree and Holt, 2004) is a model inspired by relaxing rationalizability through allowing for logistic errors. Each type plays a λ-logit response to the belief that his opponents play a $\lambda \cdot \kappa$-logit response to the belief their opponents play a $\lambda \cdot \kappa^2$-logit response to their belief, and so on, using $\kappa \in [0,1]$. The model contains quantal response equilibrium and level-1 as special cases, for $\kappa = 1$ and $\kappa = 0$, respectively.

Definition 6 (Noisy introspection, NI)**.** Given $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is consistent with (λ, κ)-noisy introspection if all types $t_i \in T_i$ of all players $i \in N$ choose $\sigma_i(\cdot|t_i) = Logit_{t_i}(\sigma^1_{-i}|\tilde{\pi}_i, \lambda)$ with

$$\sigma_i^k(\cdot|t_i) = Logit_{t_i}(\sigma_{-i}^{k+1}|\tilde{\pi}_i, \lambda \cdot \kappa^k) \tag{10}$$

The cognitive hierarchy model (Camerer et al., 2004) adapts the level-k model by assuming that level-k players do not play a logit response to the belief that all opponents are level $k-1$, but a logit response to the belief that the opponents are at any level $k' < k$ (including level-0). Players are assumed to have rational expectations about the relative frequencies of these levels, and overall levels are assumed to have Poisson distribution in the population. Given the Poisson distribution, let $f(k) = \Pr(level = k)$ denote the relative frequency of level k overall (given distribution parameter κ), and define the conditional probability $g(k'|k) = \Pr(level = k' \,|\, level < k)$. The level-0 strategy is uniform randomization, $\sigma^0(\cdot|t_i) = 1/|A_i|$.

Definition 7 (Cognitive hierarchy model, CHM)**.** Given $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is consistent with (λ, κ)-cognitive hierarchy if all types $t_i \in T_i$ of all players $i \in N$ choose $\sigma(\cdot|t_i) = \sum_{k \geq 0} f(k) \cdot \sigma^k(\cdot|t_i)$ with

$$\sigma^k(\cdot|t_i) = Logit_{t_i}(\tau_{-i}^{k-1}|\tilde{\pi}_i, \lambda) \qquad \text{and} \qquad \tau^{k-1}(\cdot|t_i) = \sum_{k'=0}^{k-1} g(k'|k) \cdot \sigma^{k'}(\cdot|t_i). \tag{11}$$

I use the parsimonious approach of Camerer et al. (2004) to capture the distribution of levels via Poisson also to complete the level-k model. Again, $f(k) = \Pr(level = k)$ denotes the relative frequency of level k in the population (given distribution parameter κ), and the level-0 strategy is uniform randomization, $\sigma^0(\cdot|t_i) = 1/|A_i|$.

Definition 8 (Level-k). Given a payoff structure $\tilde{\pi}_i \in \{\pi_i, \pi_i^{CRRA}, \pi_i^{Proj}, \pi_i^{Curse}\}$, a strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ is consistent with (λ, κ)-level-k if all types $t_i \in T_i$ of all players $i \in N$ choose $\sigma(\cdot|t_i) = \sum_{k \geq 0} f(k) \cdot \sigma^k(\cdot|t_i)$ with

$$\sigma^k(\cdot|t_i) = Logit_{t_i}(\sigma_{-i}^{k-1}|\tilde{\pi}_i, \lambda). \tag{12}$$

Subject heterogeneity, likelihood function and maximization The precision parameters λ and κ are bounded at zero and have independent gamma distributions, whereas the degrees of risk aversion, projection and cursedness are bounded at both 0 and 1 and have independent beta distributions. Thus, each subject is described by a parameter vector $\mathbf{p} \in \mathbf{P}$ with joint density $f()$. Using $o_s = (o_{s,t})$ to describe the observations of subject $s \in S$ at time $t \in T$, and $\sigma(o_{s,t}|f)$ as the probability of observation $o_{s,t}$ under density f, the individual likelihood given the observations o_s of subject s is

$$l_s(f|o_s) = \int_{\mathbf{P}} \prod_{t \in T} \sigma(o_{s,t}|\mathbf{p}) \cdot f(\mathbf{p}) d\mathbf{p}. \tag{13}$$

The predictions $\sigma(o_{s,t}|f)$ implicitly depend also on the underlying belief model, e.g. QRE or AQRE. The integral is evaluated by simulation, using quasi random numbers, see Train (2003) and e.g. the supplement to Bellemare et al. (2008). Aggregating across subjects, the log-likelihood of the respective model with parameter density f is

$$ll(f) = \sum_{s \in S} \log l_s(f|o_s). \tag{14}$$

QREs are computed using a homotopy method leaning on Turocy (2005). Parameters are estimated by maximizing the log-likelihood, sequentially applying two maximization algorithms. Initially, I use the robust, gradient-free NEWUOA algorithm (Powell, 2006) and I verify convergence using a Newton-Raphson algorithm. The estimates are tested by extensive cross-analysis to ensure that global maxima are found. All parameter estimates are provided as supplementary material.

References

Avery, C. and Kagel, J. (1997). Second-price auctions with asymmetric payoffs: An experimental investigation. *Journal of Economics & Management Strategy*, 6(3):573–603.

Bajari, P. and Hortacsu, A. (2005). Are structural estimates of auction models reasonable? evidence from experimental data. *Journal of Political Economy*, 113(4):703–741.

Bellemare, C., Kröger, S., and van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities. *Econometrica*, 76(4):815–839.

Biernacki, C., Celeux, G., and Govaert, G. (2000). Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(7):719–725.

Blanco, M., Engelmann, D., Koch, A. K., and Normann, H.-T. (2014). Preferences and beliefs in a sequential social dilemma: a within-subjects analysis. *Games and Economic Behavior*, 87(0):122–135.

Blanco, M., Engelmann, D., and Normann, H.-T. (2011). A within-subject analysis of other-regarding preferences. *Games and Economic Behavior*, 72(2):321–338.

Blavatskyy, P. and Pogrebna, G. (2010). Models of stochastic choice and decision theories: why both are important for analyzing decisions. *Journal of Applied Econometrics*, 25(6):963–986.

Bottom, W. P. and Paese, P. W. (1999). Judgment accuracy and the asymmetric cost of errors in distributive bargaining. *Group Decision and Negotiation*, 8(4):349–364.

Breitmoser, Y. (2012). Strategic reasoning in p-beauty contests. *Games and Economic Behavior*, 75(2):555–569.

Breitmoser, Y. (2013). Estimation of social preferences in generalized dictator games. *Economics Letters*, 121(2):192–197.

Breitmoser, Y. (2015). Knowing me, imagining you: Projection and overbidding in auctions. *MPRA Paper 62052*.

Browne, M. (2000). Cross-validation methods. *Journal of Mathematical Psychology*, 44(1):108–132.

Brunner, C., Camerer, C., and Goeree, J. K. (2011). Stationary concepts for experimental 2×2 games: Comment. *American Economic Review*, 101(2):1029–1040.

Camerer, C., Ho, T., and Chong, J. (2004). A cognitive hierarchy model of games. *Quarterly Journal of Economics*, 119(3):861–898.

Camerer, C. and Ho, T. H. (1999). Experience-weighted attraction learning in normal form games. *Econometrica*, 67(4):827–874.

Capra, C. M., Goeree, J. K., Gomez, R., and Holt, C. A. (1999). Anomalous behavior in a traveler's dilemma? *American Economic Review*, 89(3):678–690.

Choi, S., Gale, D., and Kariv, S. (2008). Sequential equilibrium in monotone games: A theory-based analysis of experimental data. *Journal of Economic Theory*, 143(1):302–330.

Costa-Gomes, M. and Crawford, V. (2006). Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96(5):1737–1768.

Costa-Gomes, M., Crawford, V., and Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 69(5):1193–1235.

Costa-Gomes, M., Crawford, V., and Iriberri, N. (2009). Comparing models of strategic thinking in Van Huyck, Battalio, and Beil's coordination games. *Journal of the European Economic Association*, 7(2-3):365–376.

Costa-Gomes, M. and Weizsäcker, G. (2008). Stated beliefs and play in normal-form games. *Review of Economic Studies*, 75(3):729–762.

Cox, J., Smith, V., and Walker, J. (1985). Experimental development of sealed-bid auction theory; calibrating controls for risk aversion. *American Economic Review*, 75(2):160–165.

Cox, J. C., Smith, V. L., and Walker, J. M. (1988). Theory and individual behavior of first-price auctions. *Journal of Risk and uncertainty*, 1(1):61–99.

Crawford, V. and Iriberri, N. (2007). Level-*k* auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica*, 75(6):1721–1770.

Dyer, D., Kagel, J., and Levin, D. (1989). Resolving uncertainty about the number of bidders in independent private-value auctions: An experimental analysis. *The RAND Journal of Economics*, 20(2):268–279.

Engelbrecht-Wiggans, R. (1989). The effect of regret on optimal bidding in auctions. *Management Science*, 35(6):685–692.

Engelbrecht-Wiggans, R. and Katok, E. (2007). Regret in auctions: Theory and evidence. *Economic Theory*, 33(1):81–101.

Engelmann, D. and Strobel, M. (2012). Deconstruction and reconstruction of an anomaly. *Games and Economic Behavior*, 76(2):678–689.

Erev, I. and Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88(4):848–881.

Eyster, E. and Rabin, M. (2005). Cursed equilibrium. *Econometrica*, pages 1623–1672.

Fey, M., McKelvey, R. D., and Palfrey, T. R. (1996). An experimental study of constant-sum centipede games. *International Journal of Game Theory*, 25(3):269–287.

Filiz-Ozbay, E. and Ozbay, E. (2007). Auctions with anticipated regret: theory and experiment. *American Economic Review*, 97(4):1407–1418.

Frederick, S. (2012). Overestimating others' willingness to pay. *Journal of Consumer Research*, 39(1):1–21.

Galinsky, A. D. and Mussweiler, T. (2001). First offers as anchors: the role of perspective-taking and negotiator focus. *Journal of personality and social psychology*, 81(4):657.

Garvin, S. and Kagel, J. (1994). Learning in common value auctions: Some initial observations. *Journal of Economic Behavior & Organization*, 25(3):351–372.

Goeree, J. and Holt, C. (2004). A model of noisy introspection. *Games and Economic Behavior*, 46(2):365–382.

Goeree, J. and Offerman, T. (2002). Efficiency in auctions with private and common values: An experimental study. *American Economic Review*, pages 625–643.

Goeree, J. K., Holt, C. A., and Laury, S. K. (2002a). Private costs and public benefits: Unraveling the effects of altruism and noisy behavior. *Journal of Public Economics*, 83(2):255–276.

Goeree, J. K., Holt, C. A., and Palfrey, T. R. (2002b). Quantal response equilibrium and overbidding in private-value auctions. *Journal of Economic Theory*, 104(1):247–272.

Goeree, J. K., Holt, C. A., and Palfrey, T. R. (2003). Risk averse behavior in generalized matching pennies games. *Games and Economic Behavior*, 45(1):97–113.

Haile, P. (2001). Auctions with resale markets: an application to us forest service timber sales. *American Economic Review*, 91(3):399–427.

Haile, P., Hortacsu, A., and Kosenok, G. (2008). On the empirical content of quantal response equilibrium. *American Economic Review*, 98(1):180–200.

Hey, J. (2005). Why we should not be silent about noise. *Experimental Economics*, 8(4):325–345.

Hey, J., Lotito, G., and Maffioletti, A. (2010). The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity. *Journal of risk and uncertainty*, 41(2):81–111.

Ho, T., Wang, X., and Camerer, C. (2008). Individual differences in ewa learning with partial payoff information. *The Economic Journal*, 118(525):37–59.

Kagel, J., Harstad, R., and Levin, D. (1987). Information impact and allocation rules in auctions with affiliated private values: A laboratory study. *Econometrica*, 55(6):1275–1304.

Kagel, J. and Levin, D. (1986). The winner's curse and public information in common value auctions. *American Economic Review*, 76(5):894–920.

Kagel, J. and Levin, D. (1993). Independent private value auctions: Bidder behaviour in first-, second-and third-price auctions with varying numbers of bidders. *The Economic Journal*, 103(419):868–879.

Kagel, J. and Levin, D. (2002). *Common value auctions and the winner's curse*. Princeton University Press.

Keane, M. and Wolpin, K. (2007). Exploring the usefulness of a nonrandom holdout sample for model validation: Welfare effects on female behavior. *International Economic Review*, 48(4):1351–1378.

Krueger, J. and Clement, R. W. (1994). The truly false consensus effect: an ineradicable and edocentric bias in social perception. *Journal of personality and social psychology*, 67(4):596–610.

Kurt, D. and Inman, J. J. (2013). Mispredicting others' valuations: Self-other difference in the context of endowment. *Journal of Consumer Research*, 40(1):78–89.

Loewenstein, G., O'Donoghue, T., and Rabin, M. (2003). Projection bias in predicting future utility. *The Quarterly Journal of Economics*, 118(4):1209–1248.

Madarász, K. (2012). Information projection: Model and applications. *Review of Economic Studies*, 79(3):961–985.

Madarász, K. (2015). Projection equilibrium: Definition and applications to social investment and persuasion. *Unpublished manuscript*.

McKelvey, R. D. and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38.

McLachlan, G. and Peel, D. (2000). *Finite Mixture Models*. Wiley series in probability and statistics.

Milgrom, P. (1989). Auctions and bidding: A primer. *The Journal of Economic Perspectives*, 3(3):3–22.

Mullen, B., Atkins, J. L., Champion, D. S., Edwards, C., Hardy, D., Story, J. E., and Vanderklok, M. (1985). The false consensus effect: A meta-analysis of 115 hypothesis tests. *Journal of Experimental Social Psychology*, 21(3):262–283.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313–1326.

Powell, M. (2006). The newuoa software for unconstrained optimization without derivatives. *Large-Scale Nonlinear Optimization*, pages 255–297.

Ross, L., Greene, D., and House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3):279–301.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464.

Shapiro, D., Shi, X., and Zillante, A. (2014). Level-k reasoning in a generalized beauty contest. *Games and Economic Behavior*, 86:308–329.

Stahl, D. O. and Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254.

Tang, F. (2003). A comparative study on learning in a normal form game experiment. *Journal of Economic Behavior and Organization*, 50(3):385–390.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge Univ Pr.

Turocy, T. (2005). A dynamic homotopy interpretation of the logistic quantal response equilibrium correspondence. *Games and Economic Behavior*, 51(2):243–263.

Turocy, T. (2008). Auction choice for ambiguity-averse sellers facing strategic uncertainty. *Games and Economic Behavior*, 62(1):155–179.

Weizsäcker, G. (2003). Ignoring the rationality of others: evidence from experimental normal-form games. *Games and Economic Behavior*, 44(1):145–171.

Wilcox, N. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In Cox, J. C. and Harrison, G. W., editors, *Risk aversion in experiments*, volume 12 of *Research in experimental economics*, pages 197–292. Emerald Group Publishing Limited.

Wilcox, N. (2011). Stochastically more risk averse: A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics*, 162(1):89–104.