Preferences and Performance in Simultaneous First-Price Auctions: A Structural Analysis*

Matthew Gentry† Tatiana Komarova‡ Pasquale Schiraldi§

October 2016

Abstract

Motivated by the empirical prevalence of simultaneous bidding across a wide range of auction markets, we develop and estimate a structural model of strategic interaction in simultaneous first-price auctions when objects are heterogeneous and bidders have preferences over combinations. We establish non-parametric identification of primitives in this model under standard exclusion restrictions, providing a basis for both estimation and testing of preferences over combinations. We then apply our model to data on Michigan Department of Transportation (MDOT) highway procurement auctions, quantifying the magnitude of cost synergies and evaluating the performance of the simultaneous first-price mechanism in the MDOT marketplace.

*We are grateful to Philip Haile, Ali Hortacsu, Ken Hendricks, Paul Klemperer, Paulo Somaini and Balazs Szentes for their comments, suggestions and insights. We also thank seminar and conference participants at the AEA Winter Meetings 2015, CEPR 2015, the NBER Winter Meetings 2016, the University of Wisconsin (Madison), the University of Zurich, the University of Leuven, Oxford University, City University, Cornell University, Universitie Paris 1, Yale University, Vanderbilt University, Duke University, Johns Hopkins University, Mannheim University, Emory University and the University of Maryland for helpful discussion. We thank Joachim Groeger and Paulo Somaini for graciously sharing their data. Financial support from ESRC Grant #ES/N000056/1 is gratefully acknowledged. An earlier draft of this paper was circulated under the title "Simultaneous First-Price Auctions with Preferences over Combinations: Identification, Estimation and Application."

†London School of Economics, m.l.gentry@lse.ac.uk

‡London School of Economics, t.komarova@lse.ac.uk

§London School of Economics and CEPR, p.schiraldi@lse.ac.uk

1 Introduction

Simultaneous bidding in multiple first-price auctions is a commonly occurring but rarely discussed phenomenon in many real-world auction markets.[1] In environments where values over combinations are non-additive in the set of objects won, bidders must account for possible combination wins at the time of bidding. This in turn substantially alters the strategic bidding problem compared to the standard first price auction with ambiguous welfare implications depending on the importance of synergies (either positive or negative) among objects. As a first step toward exploring this issue, we develop a structural model of bidding in simultaneous first-price auctions and study identification and estimation in this framework. We then apply our methodology to estimate cost synergies arising in Michigan Department of Transportation (MDOT) highway procurement auctions, using the resulting estimates to analyze revenue and efficiency performance of the simultaneous first-price mechanism in this application.[2]

To illustrate the policy questions arising in simultaneous multi-object auctions, note that given a set of L heterogeneous objects for sale, bidders i's preference structure could in principle be as complex as a complete 2^L-dimensional set of signals describing the valuations i assigns to each of the 2^L possible subsets of objects. Meanwhile, the simultaneous first-price mechanism allows bidders to submit (at most) L individual bids on the L objects being sold. Consequently, the simultaneous first-price auction format is necessarily inefficient – the "message space" (standalone bids) is insufficiently rich to allow bidders to express their true preferences. Allowing com-

[1] To underscore the prevalence of simultaneous bidding in applications, note that most widely studied first-price marketplaces in fact exhibit simultaneous bids. Concrete examples include markets for highway procurement in most US states (Jofret-Bonet and Pesendorfer 2003, Krasnokutskaya 2009, Krasnokutskaya and Seim 2004, Somaini 2013, Li and Zheng 2009, Groeger 2014, many others), snow-clearing in Montreal (Flambard and Perrigne 2006), recycling services in Japan (Kawai 2010), cleaning services in Sweden (Lunander and Lundberg 2012), oil and drilling rights in the US Outer Continental Shelf (Hendricks and Porter 1988, Hendricks, Pinkse and Porter 2003), and to a lesser extent US Forest Service timber harvesting (Lu and Perrigne 2008, Li and Zheng 2012, Li and Zhang 2010, Athey, Levin and Siera 2011, many others).

[2] This paper focuses on complementarities arising when auctions are run simultaneously. This complements the literature on potential linkages in valuations over time, e.g. Balat 2015, De Silva 2005, De Silva et al 2003, Groeger 2014, Jofre-Bonet and Pesendorfer 2003 among others.

binatorial bids might help to alleviate this "message space" problem, but need not produce an efficient allocation (see e.g. Crampton at al. 2006 for a review) and could impose substantial practical costs on both bidders and the seller (the "winner determination problem"). Hence in evaluating the relative merit of the simultaneous first-price format it is first necessary to assess the empirical magnitude of revenue and efficiency losses due to simultaneous bidding. Very little is presently known about these questions, due in part to the scarcity of methods for analyzing preferences over combinations in simultaneous auctions.

We develop a structural empirical model of bidding in simultaneous first-price auctions when objects are heterogeneous and bidders have non-additive preferences over combinations, to our knowledge the first such in the literature. We represent the total value i assigns to each combination as the sum of two components: the sum of the *standalone valuations* bidder i assigns to winning each object in the combination individually, plus a combination-specific *complementarity* (either positive or negative) capturing the incremental gain or loss i assigns to the combination as a whole. We interpret standalone valuations as private information drawn independently across bidders conditional on observables, but require incremental preferences over combinations to be *stable* in the sense that complementarities are functions of observables.[3] We find this framework natural in a variety of procurement contexts – when, for instance, non-additivity in preferences can be represented as realizations of a utility shock realized after a multiple win. Furthermore – and crucially – our framework collapses immediately to the standard separable model when complementarities are zero, supporting formal testing of this hypothesis.

Building on this framework, we make four main contributions to the literature on structural analysis of auction markets. First, we establish a new set of identification results applicable even when complementarities are non-zero. We first show that optimal behavior in this environment yields an inverse bidding system nonparametrically identified up to the unknown function describing complementarities,

[3] Note that this structure does not restrict dependence between i's standalone valuations for different objects in the market. We view this flexibility as critical, as in practice we expect i's standalone valuations to be positively correlated.

which collapses to the standard inverse bidding function of Guerre, Perrigne and Vuong (2000) when complementarities are zero. Under natural exclusion restrictions – namely, that marginal distributions of standalone valuations are invariant either to characteristics of other bidders or characteristics of other objects – we then translate this inverse bidding system into a system of linear equations in unknown bidder complementarities, with excludable variation in competition yielding non-parametric identification and excludable variation in other characteristics yielding semiparametric identification of these. We thereby provide a formal basis for structural analysis of simultaneous first-price auctions with non-additive preferences over combinations, to our knowledge the first such in the literature.

Second, we develop a three-step procedure by which to estimate primitives in our structural model. First, in Step 1, we estimate the multi-variate joint distribution of bids as a function of bidder- and auction-level characteristics. Due to the high-dimensional nature of this estimation problem, we follow several prior studies (e.g. Cantillon and Pesendorfer 2006 and Athey, Levin and Siera 2011) by employing a parametric approximation to the bid density in implementing this step. Next, in Step 2, we parametrize preferences over combinations as a function of bidder- and combination-specific covariates[4] and estimate parameters in this function by minimization of a simulated analogue to our semiparametric identification criterion. Finally, in Step 3, we map estimates derived in Step 2 through the inverse bidding system derived in Step 1 to obtain estimates of the distribution of private costs rationalizing observed bidding behavior.

Third, we apply our structural framework to analyze simultaneous bidding in Michigan Department of Transportation (MDOT) highway procurement markets. We view this market as prototypical of our target application: large numbers of projects are auctioned simultaneously (an average of 45 per letting round in our 2005-2015 sample period), more than half of bidders bid on at least two projects simultaneously (with an average of 2.7 bids per round across all bidders in the sam-

[4]In our application, combination-specific covariates might include the sum of engineer's estimates across projects in a combination, distance between projects in a combination, and indicators for whether projects in a combination are of the same type, among others.

ple), and combination and contingent bidding are explicitly forbidden. Within this marketplace, we show that factors such as size of other projects, number of bidders in other auctions, and the relative distance between projects have substantial reduced-form impacts on i's bid in auction l, a finding hard to rationalize in standard separable models. We then apply the three-step estimation algorithm described above to recover structural estimates of primitives. Our results suggest that a combination win would generate roughly 13 percent cost savings for a combination at the 95th (best) percentile in our sample, transitioning to roughly 3.5 percent cost increases for a combination at the 5th (worst) percentile, with large and / or heterogeneous projects more likely to be substitutes.

Finally, building on our structural estimates, we measure potential inefficiencies associated with the simultaneous first price auction design. Towards this end, we compare the simultaneous first-price auction used in the MDOT marketplace with a mechanism which ensures both an efficient allocation and provides a reasonable benchmark to compare procurement costs: the combinatorial proxy auction of Ausubel and Milgrom (2002).[5] As expected, this counterfactual alternative yields non-trivially lower social costs costs: our estimates suggest total social gains of approximately four percent, with relatively larger gains in lettings with larger complementarities. Interestingly, however, the majority of these social gains accrue to bidders: MDOT's expected procurement costs fall by only about 1 percent. In other words, even in the presence of substantial complementarities, and ignoring any other implementation costs, the benefits of switching to a combinatorial mechanism are (from MDOT's perspective) relatively small. This is to our knowledge the first structural comparison of the simultaneous first-price format with leading combinatorial alternatives, and in our view helps to rationalize the popularity of the simultaneous first-price format in applications.

While this is to our knowledge the first structural analysis of bidding in simul-

[5] Following Ausubel and Milgrom (2002), we assume for the purposes of this comparison that bidders truthfully report their valuations to the proxy bidder. Alternatively, one could instead consider the classic Vickery-Clarke-Groves (VCG) mechanism as an efficient benchmark. This would lead to the same allocations as the Ausubel-Milgrom proxy auction, but the VCG mechanism is known to have poor revenue performance.

taneous first-price auctions, our work builds on a small but growing structure literature analyzing combinatorial auctions.[6] Cantillon and Pesendorfer (2006) analyze combinatorial first-price sealed-bid auctions for London bus routes, using the possibility of package bidding to identify bidder preferences over combinations. In their framework, identification turns on invertibility of the Jacobian of the mapping between player i's bids and the equilibrium probability that player i wins each possible combination, which in turn allows one to invert the system of necessary first order conditions describing optimal bidding to recover combinatorial valuations in terms of combinatorial bids. In our setting, this procedure necessarily fails; by construction, we observe only L bids for up to $2^L - 1$ unknown combinatorial valuations. This represents a substantially different (and more challenging) identification problem, for which we develop a novel solution. More recently, Bajari and Fox (2013) have estimated the deterministic component of bidder valuations in FCC simultaneous ascending spectrum auctions without package bidding. They exploit the assumption that the allocation of licenses is pairwise stable in matches and use the maximum score estimator for matching game to estimate the valuation function. Finally, Kim, Olivares and Weintraub (2014) have extended the methodology of Cantillon and Pesendorfer (2006) to analyze the large-scale combinatorial auctions used in procurement of Chilean school meals.

Paralleling these structural studies, there is also a small reduced-form literature seeking to quantify the role of preferences over combinations in multi-object auctions. Ausubel, Cramton, McAfee and McMillan (1997) and Moreton and Spiller (1998) measure synergy effects in FCC spectrum auctions. Lunander and Lundberg (2012) empirically compare combinatorial and simultaneous first-price auctions in a Swedish market for internal cleaning services, finding that bidders inflate their standalone bids in combinatorial auctions relative to first-price auctions but that this does not significantly affect the procurer's final costs. De Silva (2005) and

[6] Although only tangentially related to our problem, there is also a growing empirical literature on *multi-unit auctions*, which focus on markets for homogeneous, divisible goods like electricity and treasury bills. See e.g. Fevrier, Preget, and Visser (2004); Chapman, McAdams and Paarsch (2007); Kastl (2011); Hortacsu and Puller (2008); Hortacsu and McAdams (2010) and Hortacsu (2011); Wolak (2007); and Reguant (2014).

De Silva, Jeitschko and Kosmopoulou (2005) analyze spatial synergies in Oklahoma Department of Transportation highway procurement auctions, finding that bidders winning earlier projects participate more often and bid more aggressively in subsequent nearby projects. These findings are consistent with the hypothesis of spatial synergies in procurement, motivating the structural model we consider here.

Finally, from a more theoretical perspective, there have been several studies analyzing strategic interaction in stylized models involving simultaneous first-price auctions; see for example Szentes and Rosenthal (1996) and Ghosh (2012). Gentry, Komarova, Schiraldi and Shin (2015) study existence and proprieties of equilibrium in a setting closely paralleling that studied here. There is also a substantial literature analyzing properties of various combinatorial auction mechanisms: Ausbel and Milgrom (2002), Ausbel and Cramton (2004), Cramton (1998, 2002, 2006), Krishna and Rosenthal (1996), Klemperer (2008, 2010), Milgrom (2000a, 2000b), and Rosenthal and Wang (1996), to mention just a few. Detailed surveys of this literature are given in de Vreis and Vorha (2003) and Cramton et al. (2006).[7]

The rest of this paper is organized as follows. Section 2 outlines the simultaneous bidding framework on which our structural model is built. Section 3 studies identification in this model. Section 4 describes the Michigan Department of Transportation (MDOT) highway procurement marketplace, while Section 5 presents our structural results. Section 6 counterfactually analyzes performance of the simultaneous first-price mechanism in the MDOT marketplace. Finally, Section 7 conclusions. Additional results are collected in a set of technical appendices: Appendix A collects technical proofs, Appendix B extends our framework to incorporate entry, and Appendices C and D present extended identification results.

[7]There is also a growing theoretical literature on simultaneous first-price auctions in computer science; see Feldman et al. 2012, and Syrgkanis 2012 among others. This literature focuses primarily on deriving bounds on the "Bayesian price of anarchy," or fractional efficiency loss, in simultaneous first-price auction markets. Positive results in this literature are largely restricted to settings with negative complementarities, and even in these settings bounds tend to be wide (e.g. Feldman et al. (2012) show that Bayesian Nash equilibrium captures at least half of total social surplus).

2 The simultaneous first-price bidding game

A set $\mathcal{N} = \{1, ..., N\}$ of risk-neutral bidders compete for (subsets of) a set $\mathcal{L} = \{1, ..., L\}$ of objects allocated via separate but simultaneous first-price auctions. For each bidder $i \in \mathcal{N}$, let $\mathcal{M}_i \subset \mathcal{L}$ be the set of auctions in which i is participating, and let $M_i \equiv \#\mathcal{M}_i$ be the number of auctions in this set.

Our analysis focuses on the bidding game arising after participation sets $\mathcal{M} \equiv (\mathcal{M}_1, ..., \mathcal{M}_N)$ are determined. We see this focus as natural for at least two reasons. First, bid-stage identification is a necessary prerequisite for studying entry; one cannot understand participation if one does not understand bidding. Second, insofar as our primary purpose is to estimate preferences over empirically relevant combinations, it is sufficient to focus on bidding taking participation as given. For completeness, Appendix B gives one example of an entry game which formally justifying our key identifying restrictions.

Combinatorial valuations In principle, bidder i may have distinct preferences over every possible combination (subset of objects) in her participation set $\mathcal{M}_i$. Following Cantillon and Pesendorfer (2006), we assume that these combinatorial preferences are described via a $2^{M_i} \times 1$ vector of *combinatorial valuations* Y_i, drawn privately by bidder i from a joint distribution $F_{Y,i}^{\mathcal{M}}$ satisfying the following properties:

Assumption 1 (Independent Private Values)**.** *Under each participation structure $\mathcal{M}$, bidder i draws private type Y_i is drawn from a continuous c.d.f. $F_{Y,i}^{\mathcal{M}}$ with support on a compact, convex set $\mathcal{Y}_i^{\mathcal{M}} \subset R^{2^{M_i}}$. Furthermore, $Y_i^0 = 0$, $F_{Y,i}^{\mathcal{M}}$ is common knowledge, and type draws are independent across bidders: $Y_i \perp Y_j$ for all $i, j \in \mathcal{N}$.*

Note that (for the moment) we consider a fully general framework allowing arbitrary combinatorial preferences; this simplifies development of the key notation and definitions we describe below. In Section 3, we specialize this to the case of *deterministic complementarities* on which our empirical analysis is based.

The bidding game Each bidder $i \in \mathcal{N}$ submits a single bid b_{il} for each auction l in her participation set $\mathcal{M}_i$. Bids are binding and bidders may not submit combinations bids. Bidding is simultaneous and objects are awarded auction by auction: for each $l \in \mathcal{L}$ the high bidder in auction l wins object l and pays her bid in auction l. For the moment, we assume that ties are broken independently across bidders and auctions; we return to this issue when discussing equilibrium below.

Let $\mathcal{B}_\ell \subset \mathbb{R}^+$ denote the set of feasible bids in auction $\ell = 1, ..., L$; without loss of generality, we take this to be a compact set. A *bid* b_i for player i is an $M_i \times 1$ vector such that $b_{i\ell} \in \mathcal{B}_\ell$ for all $\ell \in \mathcal{M}_i$. Let $\mathcal{B}_i^{\mathcal{M}} = \times_{\ell \in \mathcal{M}_i} \mathcal{B}_\ell$ denote i's action space under participation structure $\mathcal{M}$. A *distributional strategy* for bidder i in the sense of Milgrom and Weber (1985) is a measure $\sigma_i^{\mathcal{M}}$ over $\mathcal{Y}_i^{\mathcal{M}} \times \mathcal{B}_i^{\mathcal{M}}$ whose marginal over $\mathcal{Y}_i^{\mathcal{M}}$ is $\mathcal{F}_{Y,i}^{\mathcal{M}}$. Let $\sigma^{\mathcal{M}} = (\sigma_1^{\mathcal{M}}, ..., \sigma_N^{\mathcal{M}})$ denote a distributional strategy profile for all bidders, and σ_{-i}^{M} denote a distributional strategy profile for rivals of bidder i.

For notational convenience, we omit the superscript $\mathcal{M}$ for the remainder of this section. All objects defined below should be interpreted with reference to a given participation structure $\mathcal{M}$.

Outcomes Define an *outcome* ω from the perspective of bidder i as an $1 \times M_i$ vector such that for each $\ell \in \{1, ..., M_i\}$ the element $\omega_\ell = 1$ if the ℓth element of $\mathcal{M}_i$ is allocated to i and $\omega_\ell = 0$ otherwise. Similarly, let the *outcome matrix* Ω_i for bidder i be the $2^{M_i} \times M_i$ matrix whose rows describe all outcomes possible for i. For example, if if $M_i = 2$, then the (transpose of) Ω_i would be:

$$\Omega_i^T = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

With slight abuse of notation, let Y_i^ω denote the combinatorial valuation i assigns to the combination corresponding to outcome ω. Note that taking bid $b_i \in \mathcal{B}_i$ as given, i's net payoffs over possible outcomes are described by the $2^{M_i} \times 1$ vector $Y_i - \Omega_i b_i$.

Standalone valuations and complementarities Let i's *standalone valuation* for object $l \in \mathcal{M}_i$, denoted $V_{i\ell}$, be the valuation i assigns to the outcome "i wins object ℓ alone": $V_{i\ell} \equiv Y_i^{\{\ell\}}$. Similarly, let i's *standalone valuation vector*, denoted V_i, be the $M_i \times 1$ vector describing i's standalone valuations for each of the M_i objects for which i is competing. We define the *complementarity* (positive or negative) which i associates with outcome ω, denoted K_i^ω, as the difference between the combinatorial valuation i assigns to outcome ω and the sum of i's standalone valuations for objects won under ω:

$$K_i^\omega = Y_i^\omega - \omega^T V_i.$$

Vectorizing this definition yields the $2^{M_i} \times 1$ *complementarity vector*:

$$K_i \equiv Y_i - \Omega_i V_i.$$

Intuitively, $\Omega_i V_i$ describes the additive part of bidder i's preferences over combinations, while K_i describes non-additivities in i's preferences. In particular, our model reduces to the standard additively separable case if and only if $K_i = 0$.

Joint and marginal win probabilities Taking rival strategies σ_{-i} as given, let $P_i(b_i; \sigma_{-i})$ be the $2^{M_i} \times 1$ vector describing the probability distribution over outcomes arising when i submits bid b_i, with element $P_i^\omega(b_i; \sigma_{-i})$ of $P_i(b_i; \sigma_{-i})$ denoting the probability that i wins the *combination* associated with outcome ω. Similarly, let $\Gamma_i(b_i; \sigma_{-i})$ be the $M_i \times 1$ vector whose ℓth element $\Gamma_{im}(b_i; \sigma_{-i})$ describes the *marginal* probability that bidder i wins her ℓth auction taking own bid vector b_i and rival strategies σ_{-i} as given. Observe that $\Gamma_i(b_i; \sigma_{-i})$ is related to $P_i(b_i; \sigma_{-i})$ by

$$\Gamma_i(b_i; \sigma_{-i}) = \Omega_i^T P_i(b_i; \sigma_{-i}).$$

Note that if ties occur with probability zero at b_{im}, then $\Gamma_{i\ell}(b_i; \sigma_{-i})$ is simply the c.d.f. of the maximum rival bid in auction ℓ, evaluated at $b_{i\ell}$.

Interim payoffs Finally, consider bidder i with type $Y_i \in \mathcal{Y}_i$ competing against rivals who bid according to strategy profile σ_{-i}. Applying the definitions above, we can then write bidder i's interim payoff function as follows:

$$\begin{aligned}\pi_i(b_i; Y_i, \sigma_{-i}) &= (Y_i - \Omega_i b_i)^T P_i(b_i; \sigma_{-i}) \\ &= (\Omega V_i - \Omega_i b_i)^T P_i(b_i; \sigma_{-i}) + K_i^T P_i(b_i; \sigma_{-i}) \\ &= (V_i - b_i)^T \Gamma_i(b_i; \sigma_{-i}) + K_i^T P_i(b_i; \sigma_{-i}),\end{aligned} \tag{1}$$

Note that if i's preferences over combinations are additive (i.e. if $K_i = 0$), then (1) reduces to the standard separable form

$$\pi_i(b_i; v_i, \sigma_{-i}) = \sum_{m=1}^{M_i} (V_{im} - b_{im}) \Gamma_{im}(b_{im}; \sigma_{-i}).$$

In this case, standard first-price theory applied auction by auction will characterize an equilibrium of the overall simultaneous first price auction game.

3 Identification

Consider a population of simultaneous first-price lettings. In each letting t, the auctioneer offers L_t objects for auction to N_t bidders active in the marketplace (though as above not all bidders need be active in all auctions). Each bidder i then submits a vector b_{it} of sealed bids for each auction in which she is active, with bids submitted simultaneously and the high bidder in each auction winning that object.

For each bidder i and letting t, the econometrician observes bidder i's participation set $\mathcal{M}_{it}$, bidder i's bid vector b_{it}, and a vector of bidder-specific characteristics Z_{it}. To simplify notation, we will adopt the convention that Z_{it} includes $\mathcal{M}_{it}$.[8] Let $Z_t = (Z_{1t}, ..., Z_{Nt})$ describe characteristics of all bidders active in letting t.

On the auction side, the econometrician observes two sets of covariates: a vector of generic auction characteristics X_{lt} for each object l auctioned in letting t, and

[8]As above, we do not explicitly model determination of $\mathcal{M}_t$, but rather focus on bidding behavior taking realizations of $\mathcal{M}_t$ as given; we simply require $\mathcal{M}_t$ to be drawn jointly with other observables from a stable underlying process. We describe in Appendix B how our analysis can be extended to accommodate endogenous determination of $\mathcal{M}_t$ in a fully specified entry game.

(optionally) a vector of *combination characteristics* W_t taken to affect project complementarities without influencing standalone valuations. In a highway procurement context, X_{lt} would typically include factors like project size, project location, and type of work in project l, whereas W_t might include distance between projects, sum or product of project sizes, and degree of overlap in project schedules. For future reference, let $X_t \equiv (X_{1,t}, ..., X_{L_t,t})$ describe characteristics of all auctions at time t, with X_t^i and W_t^i denoting the subvectors of X_t and W_t relevant for bidder i.

Building on the first-order approach of Guerre, Perrigne and Vuong (2000), our identification analysis aims to leverage necessary conditions for best-response behavior in simultaneous first-price auctions. For analysis based on these conditions to proceed, we require the following hypotheses on bidder behavior:

Assumption 2. *For each realization (Z, W, X) in the support of (Z_t, W_t, X_t), the distribution of bids observed at (Z, W, X) are generated by play of a Bayesian Nash equilibrium. Furthermore, holding (Z, W, X) constant, the same equilibrium is played.*

We emphasize that both aspects of this statement are assumptions – formal equilibrium analysis at the level of generality we consider here would represent a fundamental breakthrough in its own right and as such is beyond the scope of this paper.[9] In this respect we parallel many prior studies on complex auction games, in which either existence (Fox and Bajari (2013) on spectrum auctions, Ausubel and Milgrom (2002) on proxy auctions) or uniqueness (Jofre-Bonet and Pesendorfer (2003), Roberts and Sweeting (2014), Somaini (2014) and references cited therein) cannot be guaranteed in general. From an applied perspective, we view Assumption 2 as natural: if $K_i = 0$, then existence is immediate and uniqueness follows under regu-

[9] "Fundamental" in the sense that existing theoretical tools appear insufficient to support such an analysis. As in multi-unit auctions, the presence of both multidimensional bids and multidimensional types leads to failure of classical differential-equations approaches to equilibrium analysis. Monotonicity-based methods widely used in multi-unit auctions – e.g. Athey (2001), McAdams (2006), and Reny (2011) – can be applied in special cases, but (due to potential failure of monotonicity) do not apply at the level of generality we consider here. Other approaches – e.g. that of Jackson, Simon, Swinkels and Zame (2002) applied in Cantillon and Pesendorfer (2006) – deliver generalizations of Bayes-Nash equilibria, but not Bayes-Nash equilibrium itself. See Gentry et al (2016) for a detailed discussion of the challenges associated with equilibrium analysis in simultaneous first-price auctions, plus results on equilibria in some special cases.

larity conditions (Lebrun 1999); otherwise, any model under which one can dispense with Assumption 2 will be misspecified. In this sense, our analysis formally embeds the classical additively separable model within a much more general (but far more challenging) framework permitting arbitrary complementarities.[10]

To leverage necessary conditions for optimal behavior, we require only the hypotheses on equilibrium behavior stated in Assumption 2. For such an analysis to yield point (rather than partial) identification of primitives, we require the equilibrium played to satisfy the following additional regularity conditions:

Assumption 3. *For each realization* (Z, W, X) *in the support of* (Z_t, W_t, X_t)*, the joint cumulative distribution function of bids for each bidder at* (Z, W, X) *is absolutely continuous, and for any auction* $l \in \mathcal{L}_t$ *and any bidders* i, j *active in auction* l*, the marginal distributions of bids* b_{il}*,* b_{jl} *at* (Z, W, X) *have common support.*

As above, under the null of separability ($K_i = 0$), these properties follow immediately from standard regularity conditions; when $K_i \neq 0$, we require them as assumptions. In practice, the main role of Assumption 3 is to ensure that marginal bid distributions are atomless, which in turn permits extension of the Guerre, Perrigne and Vuong (2000) first-order approach to identification analysis to settings with simultaneous auctions. In Appendix C, we show how to extend the analysis below to accommodate violations of Assumption 3. The main ideas of this extension closely parallel those in the text, although only yielding partial identification of model primitives.[11]

[10]Although existence in continuous bid spaces is beyond the scope of present theory, existence in any discrete bid space follows immediately from results in Milgrom and Weber (1985). Although it is conventional to interpret bid spaces as "approximately continuous," in practice almost every real-world bid space is ultimately discrete. In this sense, we see existence as a concern of more theoretical than practical importance. Appendix C provides an alternative set of partial identification results applicable in settings where discreteness is viewed as economically important.

[11]As noted when discussing Assumption 2 above, in real-world applications bid spaces are virtually always discrete. From a theoretical perspective this can be used to guarantee that equilibrium always exists; from a practical perspective it raises the question of how to deal with discreteness. Our analysis follows the literature's overwhelming convention of interpreting bid spaces as "approximately continuous." This seems natural in our application as bid increments are tiny (cents) relative to bids (thousands or millions of dollars). In settings where discreteness is perceived as economically important, one could fall back on the more general results in Appendix C.

3.1 Deterministic complementarities

Even cursory analysis of the simultaneous first-price bidding game suggests a major empirical challenge: bidder i's private type Y_i could involve up to $2^{M_i}-1$ distinct combinatorial valuations, but we observe only M_i bids corresponding to these $2^{M_i}-1$ unknowns. To make empirical progress, it is therefore necessary to impose additional structure. We therefore propose to specialize the empirical model as follows: whereas standalone valuations (V_i) remain private information as in the standard separable model, complementarities K_i are stable functions of bidder, auction and combination specific unobservables. We formalize this assumption as follows:

Assumption 4 (Stochastic V_i, stable K_i)**.** *For all i, standalone valutations V_i are distributed according to joint c.d.f. $F_i(\cdot|Z,W,X)$ but complementarities $K_i = \kappa_i(Z,W,X)$, with $V_i \perp V_j$ for all i and j and both F_i and $\kappa_i(\cdot)$ common knowledge.*

We view this structure as natural in applications such as procurement contracting for at least two reasons. First, Assumption 4 reflects our interpretation of K_i^ω as a pure combination effect; i.e. an incremental cost or benefit derived from winning objects in ω together, or the expectation over a combination-specific shock realized after a multiple win. Second, as above, Assumption 4 naturally nests the null hypothesis of additively separable preferences: $\kappa_i(Z,W,X)=0$ for all (X,W,Z). It therefore provides an ideal framework within which to evaluate this hypothesis.

As usual, our identification analysis proceeds holding generic auction characteristics X fixed. For simplicity, we therefore omit X in notation in this section. All statements that follow should be interpreted as applying pointwise in X.

3.2 Nonparametric identification up to κ_i

Holding X constant, let $G_i(\cdot|Z,W)$ be the c.d.f. of the equilibrium joint distribution of the $M_i \times 1$ bid vector b_i submitted by bidder i at observables (Z,W,X). Let $P_{-i}(\cdot|Z,W) : \mathcal{B}_i \to \Delta^{2_i^M}$ be the probability distribution over outcomes facing bidder i taking rival strategies at (Z,W) as given, and $\Gamma_{-i}(\cdot|Z,W) \equiv \Omega^T P_{-i}(\cdot|Z,W)$ be the $M_i \times 1$ vector of marginal win probabilities corresponding to $P_{-i}(\cdot|Z,W)$.

Note that under Assumption 2, $G_i(\cdot|Z,W)$ is identified directly from the data for all i, with identification of $G_1, ..., G_N$ implying identification of P_{-i} and Γ_{-i}. Given any realization v_i of V_i and any candidate complementarity vector K_i, we can therefore write the problem facing bidder i at observables (Z,W) in terms of directly identified objects as follows:

$$\max_{b\in\mathcal{B}_i}\{(v_i - b)\cdot\Gamma_{-i}(b|Z,W) + P_{-i}(b|W,Z)^T K_i\}.$$

Let $\mathcal{K}_i$ be the $(2^{M_i} - M_i - 1)$-dimensional subspace of $\mathbb{R}^{2^{M_i}}$ containing all 2^{M_i}-dimensional vectors whose first $M_i + 1$ components are zero:[12]

$$\mathcal{K}_i = \{k \in \mathbb{R}^{2^{M_i}} \,:\, k_1 = k_2 = \ldots = k_{M_i+1} = 0\}.$$

Temporarily suppose that i's objective is differentiable at $b_i \in \text{int}(\mathcal{B}_i)$; Proposition 1 establishes that under Assumption 3 this holds almost surely with respect to the measure on $\mathcal{B}_i$ induced by G_i. Then by hypothesis of equilibrium play, b_i must satisfy necessary first-order conditions for an interior optimum:

$$\nabla_b\Gamma_{-i}(b_i|Z,W)(v_i - b^*) = \Gamma_{-i}(b_i|Z,W) - \nabla_b P_{-i}(b_i|W,Z)^T K_i. \tag{2}$$

Clearly, this system is not invertible for (v_i, K_i) jointly. But *taking* $K_i \in \mathcal{K}_i$ *as given*, it reduces to an $M_i \times 1$ system of equations in the $M_i \times 1$ vector of unknown standalone valuations v_i. Under our assumptions, this system *is* invertible, implying that for any *candidate* $K_i \in \mathcal{K}_i$ there exists a unique candidate for v_i at which b_i satisfies first order necessary conditions for a best response:[13]

Proposition 1 (Inverse Bidding Function)**.** *Suppose that Assumptions 1-3 hold. Let $K_i \in \mathcal{K}_i$ be any candidate for bidder i's complementarity vector $\kappa_i(Z,W)$. Then for almost every (Z,W) and almost every b_i drawn from $G_i(\cdot|Z,W)$, there exists a*

[12] These zero components correspond to the outcomes in which bidder i wins either no objects ($\omega = (0,\ldots,0)$) or one object ($\omega'\omega = 1$), for which complementarities are zero by construction.

[13] Obviously, imposing *sufficient* conditions for b_i to be a best response – by, for instance, requiring second-order conditions to hold at $\xi(b_i|K_i;Z,W)$ – can only improve identification.

unique vector $\tilde{v}_i \in \mathbb{R}^{M_i}$ *satisfying the first-order system* (2) *at* b_i *under the hypothesis* $\kappa_i(Z, W) = K_i$. *This* $\tilde{v}_i$ *can be expressed in terms of* b_i *as*

$$\tilde{v}_i = \xi_i(b_i|K_i; Z, W),$$

where $\xi_i(\cdot|\cdot; Z, W) : \mathcal{B}_i \times \mathcal{K}_i \to \mathbb{R}^{M_i}$ *is defined by*

$$\begin{aligned} \xi_i(b|K_i; Z, W) \equiv \; & b + \nabla_b \Gamma_{-i}(b|Z, W)^{-1} \times \Gamma_{-i}(b|Z, W) \\ & - \nabla_b \Gamma_{-i}(b|Z, W)^{-1} \times \nabla_b P_{-i}(b|Z, W)^T K_i, \end{aligned} \quad (3)$$

and the right-hand expression is identified up to K_i.

Proof. See Appendix A. □

Recall that $\Gamma(b_i; Z, W)$ is an $M_i \times 1$ vector whose mth element is the c.d.f. of the maximum rival bid in the mth auction played by i, in which case $\nabla_b \Gamma(b_i; Z, W)$ will be a diagonal matrix whose mth diagonal element describes the corresponding rival bid density. Hence if $K_i = 0$, then $\xi_i(\cdot)$ reduces to the standard inverse bidding function of Guerre, Perrigne and Vuong (2000) defined auction-by-auction. Proposition 1 simply extends this observation to arbitrary $K_i \in \mathcal{K}_i$.

Finally, observe that if the conjecture $\kappa_i(Z, W) = K_i$ is in fact correct, then by Proposition 1 we must have $v_i = \xi_i(b_i|K_i; Z, W)$ almost surely. Hence for each candidate $K_i \in \mathcal{K}_i$, there exists a *unique identified* candidate $\tilde{F}_i(\cdot|K_i; Z, W)$ for the unknown c.d.f. $F_i(\cdot|Z, W)$ consistent with the hypothesis $\kappa_i(Z, W) = K_i$:

$$\tilde{F}_i(v|K_i; Z, W) = \int_{\mathbf{B}_i} 1[\xi_i(B_i|K_i; Z, W) \leq v] \, G_i(dB_i|Z, W). \quad (4)$$

Since $F_i(\cdot|Z, W) \equiv \tilde{F}_i(\cdot|\kappa_i(Z, W); Z, W)$ by construction, it follows that identification of the model reduces to identification of $\kappa_i(Z, W)$. We thus now turn to consider identification of κ_i, both non-parametrically through variation in rival characteristics Z_{-i} and semi-parametrically through variation in combination characteristics W.

3.3 Nonparametric identification of κ_i with excluded competitor characteristics

Now suppose that to the assumptions described so far, we add the restriction that bidder i's primitives F_i, κ_i depend only on bidder i's type Z_i:

Assumption 5. *$F_i(\cdot|Z, W) = F_i(\cdot|Z_i, W)$ and $\kappa_i(Z, W) = \kappa_i(Z_i, W)$.*

Similar assumptions have been widely employed in the empirical auction literature; see, e.g., Haile, Hong and Shum (2003), Guerre, Perrigne and Vuong (2009), and Somaini (2014) among others. We will show that under Assumption 5, variation in competitor characteristics Z_{-i} induces a large (infinite) set of restrictions on the finite vector $\kappa_i(Z_i, W)$. Under mild conditions on variation in Z_{-i} made precise below, this system will have the unique (overdetermined) solution $\kappa_i(Z_i, W) \in \mathcal{K}_i$, leading to nonparametric identification of $\kappa_i(Z_i, W)$ (and hence the model as above).

To understand how variation in Z_{-i} identifies $\kappa_i(Z_i, W)$, consider a simple two-auction example. Starting from some initial set of competitor characteristics Z_{-i}, let Z'_{-i} be the competitor characteristics derived from Z_i by adding, for example, one additional bidder to Auction 2. Then the *marginal* probability that i wins Auction 1 will be similar at Z_{-i} and Z'_{-i}, but the probability of the joint outcome "i wins both 1 and 2" will differ. Furthermore, under Assumption 5, this is the *only* way that shifting Z'_{-i} matters in Auction 1. Therefore to the extent that moving from Z_{-i} to Z'_{-i} matters for i's behavior in Auction 1, it can be only through $\kappa_i(Z_i, W)$; in particular, if $\kappa_i(Z_i, W) = 0$, then we should see no effect. Since the set of such "experiments" is limited only by the support of $Z_{-i}|Z_i, W$, this in turn provides a powerful source of identifying information on $\kappa_i(Z_i, W)$.

We now formalize this intuition. By linearity of $\xi_i(\cdot|K_i; Z, W)$ in K_i, we have for any (Z, W) and any $K \in \mathcal{K}_i$:

$$E_{B_i}[\xi_i(B_i|K; Z, W) \,|\, Z, W] \equiv \Upsilon_i(Z, W) - \Psi_i(Z, W) \cdot K, \tag{5}$$

where $\Upsilon_i(Z,W)$ is an identified $M_i \times 1$ vector defined by

$$\Upsilon_i(Z,W) = \int_{\mathcal{B}_i} \left(B_i + \nabla_b \Gamma_{-i}(B_i|Z,W)^{-1} \Gamma_{-i}(B_i|Z,W) \right) G_i(dB_i|Z,W)$$

and $\Psi_i(Z,W)$ is an identified $M_i \times 2^{M_i}$ matrix defined by

$$\Psi_i(Z,W) = \int_{\mathcal{B}_i} \nabla_b \Gamma_{-i}(B_i|Z,W)^{-1} \nabla_b P_{-i}(B_i|Z,W)^T \, G_i(dB_i|Z,W).$$

Now consider any Z_i and any Z_{-i}, Z'_{-i} in the support of $Z_{-i}|Z_i, W$. From above, we have $F_i(\cdot|Z_i,W) \equiv \tilde{F}_i(\cdot|\kappa_i(Z_i,W); Z, W)$ for all Z, hence in particular

$$\tilde{F}_i(\cdot|\kappa_i(Z_i,W); Z_i, Z_{-i}, W) \equiv \tilde{F}_i(\cdot|\kappa_i(Z_i,W); Z_i, Z'_{-i}, W). \tag{6}$$

But invariance of distributions implies invariance of expectations, hence letting $Z = (Z_i, Z_{-i})$ and $Z' = (Z_i, Z'_{-i})$:

$$E_{B_i}[\,\xi(B_i|\kappa_i(Z_i,W); Z, W) \,|\, Z, W] = E_{B_i}[\xi(B_i|\kappa_i(Z_i,W); Z', W)|Z', W]. \tag{7}$$

Substituting (5) into (7), we obtain an $M_i \times 1$ system of linear equations in the $2^{M_i} \times 1$ vector of unknowns $\kappa(Z_i, W) \in \mathcal{K}_i$:

$$[\Upsilon_i(Z,W) - \Upsilon_i(Z',W)] - [\Psi_i(Z,W) - \Psi_i(Z',W)] \cdot \kappa(Z_i,W) = 0. \tag{8}$$

For a single Z_{-i}, Z'_{-i} pair, this system will typically be rank-deficient. But the underlying equality restriction must hold for *every* Z_{-i}, Z'_{-i} in the support of $Z_{-i}|Z_i, W$. Pooling these restrictions across Z_{-i}, Z'_{-i}, we ultimately conclude:

Proposition 2. *Suppose there exist vectors $Z_{-i,0}, Z_{-i,1}, ..., Z_{-i,J}$ in the support of $Z_{-i}|Z_i, W$ such that the submatrix $\overline{\mathbf{M}}_Z$ formed by that last $(2^{M_i} - M_i - 1)$ columns* *of the* $JM_i \times 2^{M_i}$ *matrix*

$$\mathbf{M}_Z \equiv \begin{bmatrix} \Psi_i(Z_i, Z_{-i,1}, W) - \Psi_i(Z_i, Z_{-i,0}, W) \\ \vdots \\ \Psi_i(Z_i, Z_{-i,J}, W) - \Psi_i(Z_i, Z_{-i,0}, W) \end{bmatrix}$$

has rank $2^{M_i} - M_i - 1$*. Then* $\kappa_i(Z_i, W)$ *is identified.*

Recall that the identification criterion (8) exploits only invariance of *first moments* of $F_i(\cdot|Z_i, W)$ across competitor characteristics Z_{-i}, whereas the underlying distributional invariance restriction (6) requires equality of *all moments*. The system of equations in Proposition 2 merely provides a simple and testable sufficient condition under which the underlying system has a unique solution. Note also that variation in, e.g., number of rivals in each auctions will produce exactly the kind of variation needed for full column rank of $\overline{\mathbf{M}}_Z$: nonlinear changes in combination win probabilities which matter for cross-auction bidding only through $\kappa_i(Z_i, W)$. Even discrete variation in Z_{-i} thus naturally gives rise to full column rank of $\overline{\mathbf{M}}_\Psi$, yielding nonparametric identification of primitives as above.

3.4 Semiparametric identification of κ_i with excludable combination characteristics

While the restriction that own primitives are invariant to competitor characteristics is both natural and widely employed, it could potentially be violated in environments with richer strategic interaction among players. For instance, if there is a competitive upstream market for sub-contractors, then capacity utilization by i's rivals could in principle affect i's costs. We therefore also consider semi-parametric identification of $\kappa_i(\cdot)$ based on excludable variation in combination characteristics W. In this approach, we replace Assumption 5 with two alternative identifying assumptions: a parametric form for $\kappa_i(Z, W)$ and an exclusion restriction on W:

Assumption 6 (κ_i linear in parameters)**.** $\kappa_i(Z, W) = \mathbf{C}_i(Z_i, W) \cdot \theta_{0i}$*, where* $\mathbf{C}_i(Z_i, W)$ *is a known mapping from* (Z_i, W_i) *to* $\mathbb{R}^{p_i}$ *and* $\theta_{0i} \in \Theta_i$ *is an unknown* $p_i \times 1$ *vector* *of parameters.*

Assumption 7 (Standalone valuations invariant to W)**.** $F_i(\cdot|Z, W) = F_i(\cdot|Z)$.

We see some version of Assumption 6 as natural since we will typically wish to parameterize κ_i in practice; the linear-in-parameters structure considered here is inessential and serves mainly to simplify the analysis. Meanwhile, Assumption 7 simply formalizes the exclusion restriction underlying the definition of W: i.e. that W contains factors which shift complementarities without shifting standalone valuations. For instance, in our application, W includes factors such as distance between projects (holding distance to the bidder constant) or overlap in project schedules (holding project length constant). For such variables we see Assumption 7 as quite natural.

Now taking $Z = (Z_i, Z_{-i})$ as given, consider identification of κ_i based on variation in W. Let $(W^0, W^1, ..., W^J)$ be any collection of realizations of W, and for each $j = 0, 1, ..., J$ let $\Upsilon_i^j \equiv \Upsilon_i(Z, W^j)$, $\Psi_i^j \equiv \Upsilon_i(Z, W^j)$, and $\mathbf{C}_i^j \equiv \mathbf{C}_i(Z_i, W^j)$ denote values of the functions $\Upsilon_i(\cdot)$, $\Psi_i(\cdot)$, and $\mathbf{C}_i(\cdot)$ evaluated at (Z, W^j). Substituting $\kappa_i(Z, W^j) \equiv \mathbf{C}_i^j \theta_{0i}$ into Equation 8, it follows that true parameters θ_{0i} must satisfy:

$$(\Upsilon_i^j - \Upsilon_i^k) = (\Psi_i^j \mathbf{C}_i^j - \Psi_i^k \mathbf{C}_i^k) \cdot \theta_{0i} \quad \forall\, j, k \in \{0, 1, ..., J\}. \tag{9}$$

While for any given (j, k) pair this system might be rank deficient, the number of available "experiments" $(W^0, W^1, ..., W^J)$ is again limited only by the support of $W|Z$. So long as variation in $W|Z$ is sufficiently rich in the following (weak) sense, it follows that the model is semiparametrically identified:

Proposition 3. *Suppose there exist vectors $W^0, W^1, ..., Z^J$ in the support of $W|Z$ such that the $JM_i \times p_i$ matrix $\mathbf{M}_W$ defined by*

$$\mathbf{M}_W \equiv \begin{bmatrix} \Psi_i^1 \mathbf{C}_i^j - \Psi_i^0 \mathbf{C}_i^0 \\ \vdots \\ \Psi_i^J \mathbf{C}_i^j - \Psi_i^0 \mathbf{C}_i^0 \end{bmatrix}$$

has rank p_i. Then θ_{0i} (and hence $F_i(\cdot|Z)$) is identified.

Note that although for simplicity we have omitted generic auction covariates X in the discussion so far, it may be that some elements of W (for instance, total size of a combination) depend deterministically on X. In this case, as in our application, we proceed instead under the following variant of Assumption 7:

Assumption 8. *For all bidders i and auctions l, $F_{il}(\cdot|Z, W, X) = F_{il}(\cdot|Z, X_l)$.*

This assumption imposes two key restrictions: the *marginal* distribution of each standalone valuation V_{il} is invariant to W, and (conditional on Z) this marginal distribution is invariant to characteristics in auctions other than l.[14] The identification argument changes in only two respects: the invariance condition (9) is evaluated separately for each V_{il} (rather than jointly over V_i), and identification arises from variation in $W|Z, X_l$ rather than variation in $W|Z, X$ as above.

4 Application: Michigan Highway Procurement

We now turn to our empirical application: the marketplace for Michigan Department of Transportation (MDOT) highway construction and maintenance contracts. As common in similar procurement contexts, MDOT allocates contracts for a wide range of highway construction and maintenance services via low-price sealed-bid auctions. The vast majority of MDOT projects are allocated via large simultaneous letting rounds, which take place on average every three weeks.[15] There are an average of 45 auctions per letting round and more than half (56 percent) of bidders submit bids on multiple contracts in any given letting.[16] A bid is an itemized description of unit costs for each line item specified in contract plans; bids are submitted to MDOT project by project, with the winner of each project the bidder submitting the bid involving the lowest total project costs. Contracts are advertised up to

[14]Since this assumption deals with marginal rather than joint distributions, it is neither strictly stronger nor strictly weaker than Assumption 7.

[15]There are only two months without lettings.

[16]MDOT runs a pre-qualification process, which ensures quality of work. The process involves a check on the financial status of the firm and its backlogs from all construction activities. A bid submission includes a detailed break down of all costs involved in the contract. The winner is determined solely by the total cost of the project.

ten weeks prior to letting, with the closing deadline for submitting, amending or withdrawing bids typically 10am on the letting date. MDOT then publically opens bids and allocates contracts, with opened bids legally binding and winning bidders held liable to complete contracts won. In view of prior work documenting proximity effects and capacity constraints in highway procurement, we expect factors such as capacity constraints, project proximity, project types, and scheduling overlap to induce substantial non-additivities in bidder payoffs across auctions.

4.1 Data

MDOT provides detailed records on contracts auctioned, bids received, and letting outcomes on its letting website (http://www.michigan.gov/mdot). Drawing from these records, we observe data on (almost) all contracts auctioned by MDOT over the sample period January 2005 to March 2014.[17] Our sample includes a total of 8224 auctions, where for each auction the following information is observed: project description, project location, pre-qualification requirements, the internal MDOT engineer's estimate of total project cost, and the list of participating firms and their bids. Based on project descriptions, we classify projects into five project types: bridge work, major construction, paving (primarily hot-mix asphalt), safety (e.g. signing and signals), and miscellaneous, leading to a final distribution of projects across types summarized in Table 1. As evident from Table 1, roughly 80 percent of contracts are for road and bridge construction and maintenance broadly defined, with the remainder split between safety and other miscellaneous construction.

The data contains information on a total of 859 unique bidders active in the MDOT marketplace over our sample period, which we subclassify by size and scope of activity as follows. We define "regular" bidders to be those who have submitted more than 100 bids in the sample period. This yields a total of 36 regular bidders in our sample, with all remaining bidders classified as "fringe." For the subsample of bidders who have submitted more than 50 bids, we also collect data on number and

[17]MDOT records for a small number of contracts are incomplete. Although we have data from October 2002 to March 2014, we have discarded the first few years (from October 2002 to December 2004) so to construct bidder backlog variables.

Table 1: Summary of Projects by Type

Contract Type	**Frequency**
Bridge	13.33
Major Construction	9.64
Paving	56.33
Safety	12.25
Miscellaneous	8.45

location of plants by firm. This data is derived from a variety of sources: OneSource North America Business Browser, Dun and Bradstreet, Hoover's, Yellowpages.com and firms' websites. Based on this information, we further classify bidders as "large" or "small" based on their number of plants, with "large" bidders defined as those with at least 5 plants in Michigan. We thus obtain a final classification of 8 large regular bidders, 28 small regular bidders, and 823 fringe bidders (of which 4 are also large bidders) in the MDOT marketplace.

Table 2 surveys the auction side of the MDOT marketplace. The first key feature emerging from this table is, not surprisingly, the large number of contracts auctioned simultaneously in the market: a mean of 45 per letting, with a maximum of 133 on a single letting date (note that smaller supplemental lettings are occasionally held two or three weeks after the main letting in a given month). On average about five bids are received per contract, which is small relative to the average number of bidders (approximately 84) active in any given letting. For each contract, MDOT prepares an internal "Engineer's Estimate" of expected procurement cost released to bidders before bidding; as evident from the dispersion in this measure, projects in the marketplace vary substantially in size and complexity. The statistic "Money Left on the Table" measures the percent difference between lowest and second-lowest bids; on average this is 7.4 percent or roughly $112,000 per contract, suggesting the presence of substantial uncertainty in the marketplace.

Table 3 re-frames the auction-level variables in Table 2 to provide a clearer picture of bidder behavior in the MDOT marketplace. Again, the key pattern emerging from Table 3 is the prevalence of simultaneous bidding in MDOT auctions, with the

Table 2: Auction Level Summary Statistics

	Mean	St. Dev.	Min	Max
Auctions per Round	45.19	35.67	1	133
Total Bids per Round	228.1	180.9	1	669
Distinct Bidders per Round	83.97	57.06	1	207
Number of Bidders per Auction	5.048	3.186	1	28
Large Regular Bidders per Auction	0.398	0.672	0	3
Regular Bidders per Auction	1.500	1.362	0	7
Fringe Bidders per Auction	3.149	2.926	0	23
Engineer's Estimate (in thousands)	1,514	4,689	4.412	165,313
Project Duration (in days)	175.8	205.1	2	1,838
Money Left on the Table	0.0744	0.0966	0	3.016

Table 3: Bidder Level Summary Statistics

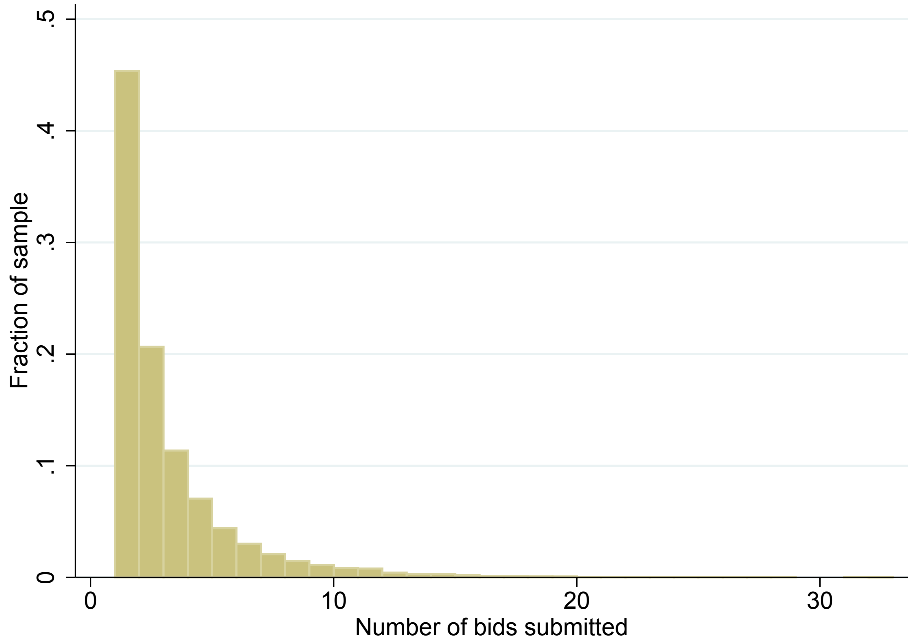
	Mean	St. Dev.	Min	Max
Bids by Round	2.716	2.785	1	33
Bids by Round if Large	6.65	6.27	1.000	33.000
Bids by Round if Regular	5.96	4.58	1.00	33.00
Backlog (in millions)	5.792	19.01	0	275.5

average bidder competing in roughly 2.7 auctions per round and large and regular bidders competing in substantially more. The variable "backlog" provides a bidder-specific measure of capacity utilization. As usual in the literature, we define backlog for bidder i at date t as the sum of work remaining among projects l won by i up to t, where work remaining on project l at date t is defined as total project size (measured by the engineer's estimate) times the proportion of scheduled project days remaining at date t. Note that number of bids submitted by any given bidder is small relative to the number of bidders in the marketplace, with even large bidders competing in less than fifteen percent of total auctions on average.

Finally, Figure 1 plots the histogram (over all bidders i and lettings t) of the number of bids submitted by bidder i in letting t.[18] As evident from Figure 1, more than 55 percent of active bidders submit multiple bids in the same letting. Despite

[18] An observation for the purposes of Figure 1 is thus a bidder-letting pair.

Figure 1: Number of Simultaneous Bids Submitted, Bidder by Letting



this, it is relatively uncommon for a typical bidder to compete in a large number of auctions; almost 90 percent of bidders in our sample bid in 6 or fewer auctions and only 2 percent bid in more than 10.

4.2 Descriptive regressions

We next explore a series of simple regressions designed to explore the potential economic implications of simultaneous bidding in the MDOT marketplace. The unit of analysis in these regressions is a bidder-auction-round combination, with the dependent variable log of bid submitted by bidder i in auction l in letting t. We regress log bids on a vector of regressors intended to capture effects of own-auction and cross-auction characteristics on i's bid in auction l at time t.

Regression specification As usual, we control for a number of auction-level characteristics which we expect to be key direct determinants of i's bid in auction l: the size of auction l, proxied by the MDOT engineer's estimate of expected project cost, the level of competition i faces in auction l, and the distance between project l and i's base of operations.[19] To explore potential cross-auction interaction in the MDOT marketplace, we seek a set of covariates relevant for bidding in auction l only through κ_i, i.e. factors shifting *combination* payoffs but irrelevant for standalone valuations after conditioning on characteristics of auction l, as specified below.

To control for cross-auction competition which may shift combination win probabilities, we consider the number of rivals across all auctions played by bidder i. The effects of cross-auction competition on i's bids in auction l are theoretically ambiguous, depending both on the sign of κ_i and on strategic responses by rival bidders. A priori, however, if objects are substitutes, we expect greater competition in auction k to increase marginal returns to winning auction l.

To capture the presence of capacity constraints or diseconomies of scale, we consider two variables. First, as a direct measure of total project size, we consider the (log of) the sum of engineer's estimates across all auctions in which i is bidding. Second, as a measure of the degree of schedule overlap on projects for which i is bidding, we consider the total number of overlapping days for projects for which i submits bids, scaled by the sum of days scheduled for each of these projects. Insofar as marginal costs are increasing in capacity utilization, we expect the coefficients on these variables to be positive.

In principle, complementarities arising between similar projects may differ from those arising between different projects. To account for this possibility, we consider the Herfindahl index for project types for which bidder i is bidding. A negative sign is interpreted as a relative complementarity between similar projects.

Finally, as an additional proxy for potential economies and / or diseconomies among projects, we compute a measure of distance between projects, defined as the (log of) distance between the current project and the other projects in which i bids normalized by the total distance between each of these projects and the closest plant

[19] We construct for each bidder-project pair the minimum straight-line distance (in miles) between any of i's plants and the centroid of the county in which project l is located. We take the shortest distance if bidder i owns multiple plants.

owned by bidder i. Insofar as relatively more distant projects potentially reduce economies of scale, We expect this variable to have a positive sign.

Regression results Table 4 reports OLS estimates for our baseline regression specifications: logs bids by bidder, round, and auction on the own- and cross-auction characteristics defined above. All regression specifications include a full set of bidder type, project type, and letting date indicators, with standard errors clustered at the bidder-round level to allow for correlation within bidder i's bids.

Estimated effects of own-auction characteristics correspond closely both to our prior and to findings elsewhere in the literature. As expected, bids are increasing almost one for one in project size, with the coefficient on log engineer's estimate exceeding 0.97 in all specifications. Similarly, the negative coefficient on number of rivals suggests that competition increases bidder aggressiveness, with one additional competitor associated with a $4-5$ percent decrease in average bids. Finally, the coefficient on log distance to project suggests that a one percent increase in i's distance from the project leads to about a 2 percent increase in i's bid on average.

More importantly, estimated cross-auction effects are also highly significant, with magnitudes stable across specifications and signs broadly consistent with our prior expectations. In particular, the positive coefficient on log of engineer's estimates across auctions suggests that competing in larger auctions leads to a substantial decrease in aggressiveness by bidder i in auction l, with the negative coefficient on same-type projects suggesting that this effect is ameliorated slightly when the two projects are of the same type. Similarly, the positive sign on log distance among projects suggests that increasing distance to other projects reduces the synergies among them, which corroborates the hypothesis that simultaneous bidding induces strategic spillovers. Finally, the significant negative coefficient on total number of rivals in auctions participated by i suggests that facing more competition across auctions leads bidder i to bid more aggressively in auction l.

Table 4: OLS Estimates of Cross-Auction Effects

$y = ln(bid)$	1	2
Log engineer's estimate	0.971***	0.9765***
	(0.0011)	(0.0011)
Log number of rivals	-0.0499***	-0.0398***
	(0.0032)	(0.003)
Log distance to project	0.021**	0.0129***
	(0.0011)	(0.001)
Log days to project start	0.0038***	0.0038***
	(0.0009)	(0.0009)
Standardized backlog	0.0029***	0.0033***
	(0.001)	(0.0011)
Log number of big rivals faced	0.0047***	0.0049**
	(0.0024)	(0.0022)
Log number of regular rivals faced	0.026***	0.0304***
	(0.0031)	(0.0028)
Multiple-bid indicator	-0.0951***	-0.1805***
	(0.023)	(0.0223)
Log sum engineer's estimate across played auctions	0.0062***	0.0119***
	(0.0016)	(0.0015)
Log sum number of rivals across played auctions	-0.0162***	-0.0124***
	(0.0025)	(0.0023)
Log distance across played projects	0.0047**	0.0042**
	(0.002)	(0.002)
Fraction overlapping time across projects	0.0175***	0.0139***
	(0.0037)	(0.004)
Same-type-auctions concentration index	-0.0101**	-0.0267***
	(0.0051)	(0.0053)
Big bidder	-	0.0019
	-	(0.0045)
Regular Bidder	-	-0.0044*
	-	(0.0026)
Year FE, Month FE, Auction type FE	YES	YES
Bidder type FE	NO	YES
Bidder ID FE	YES	NO
R-squared	98.16	97.93

Unit of analysis is bidder-auction-round, with standard errors clustered by bidder within each round. There are 40,624 observations. Variables *log of engineer's estimate*, *log of number of rivals in the auction* and *log of distance to the county centroid* measure size, strength of competition, and distance to project l respectively. Remaining variables proxy for cross-auction characteristics: number of rivals in other auctions, sum engineer's estimate, distance to auctions scaled by distance to project l in which i is competing and number of overlapping days among projects scaled by the total number of days to completion.

5 Structural estimation of complementarities

We now turn to this paper's primary interest: structural estimation of the function $\kappa_i(\cdot)$ describing preferences over combinations. In principle, the results in Section 3 support fully non-parametric estimation of κ_i. In practice, of course, the dimensionality of the problem renders this infeasible. We therefore implement our structural procedure in two steps. First, following Athey, Levin and Siera (2011) and Cantillon and Pesendorfer (2006) among others, we estimate a parametric approximation to the equilibrium distribution G_i of bids submitted by each bidder i. Second, we map these estimates through the first-order condition (2) to obtain a minimum-distance criterion paralleling Equation (7) which we use to estimate parameters in κ_i. Following Groeger (2014), we assume there is no binding reserve price. When a bidder is the sole participant (which happens only 137 times out of 8824 auction analyzed), he will face MDOT that draws a completion cost from a fringe bidder's cost.

5.1 First step: estimation of $G_1, ..., G_N$

In constructing first-step estimates of $G_1, ..., G_N$, we model i's bid in auction l and letting t as depending on the following observables: i's type, characteristics X_{ilt} influencing i's standalone valuation for contract l, characteristics W_{ilt} relevant for i's preferences over combinations involving auction l, competition in auction l, and competition in other auctions in which i bids. In particular, for bidder i at (Z_t, W_t, X_t), we specify and estimate a first-step model of the form:

$$\ln(b_{it}) \sim \text{MVN}(\cdot|\mu_{ilt}, \Sigma_{ilt}),$$

where μ_{ilt} is a linear-in-parameters function of the form

$$\mu_{ilt} = \alpha \cdot \; D_{il}^{\mu}(Z_t, W_t, X_t),$$

we specify the variance and covariance components of Σ_{ilt} respectively as

$$\sigma^2_{ilt} = \exp(\beta \cdot D^{\sigma}_{il}(Z_t, W_t, X_t)),$$
$$\rho_{ilkt} = \frac{\exp(\gamma \cdot D^{\rho}_{ikl}(Z_t, W_t, X_t) - 1)}{\exp(\gamma \cdot D^{\rho}_{ikl}(Z_t, W_t, X_t) + 1)},$$

and $D^{\mu}_{il}(\cdot)$, $D^{\sigma}_{il}(\cdot)$, and $D^{\rho}_{ikl}(\cdot)$ are known (user-specified) subvectors of (Z_t, W_t, X_t).

Table 5 reports first-step estimates from applying the first-step model above to the sample described in Section 4 once we restrict our attention (for computational reason) only to bidders who bids in up to 16 auctions which represents 96% of all bidders. Panel 1 reports estimates $\hat{\alpha}$ for the parameters α appearing in the mean function μ_{it}; not surprisingly, are very similar to those in our descriptive regressions. Panel 2 reports estimates $\hat{\beta}$ for parameters β appearing in the variance function σ^2_{ilt}, which suggest that bidders competing in multiple auctions and for larger projects submit less dispersed bids.[20] While we do not have strong priors on these effects, the direction seems natural. Finally, Panel 3 reports estimates $\hat{\gamma}$ for parameters γ appearing in the covariance function ρ_{ilkt}. These suggest at least two broad patterns in bidding behavior across auctions. First, bidders bid tend to bid more similarly for similar projects: i.e. for those in the same county and / or of the same type. Second, bidders competing for projects whose schedules overlap in time tend to compete for one relatively more aggressively than for the other. We interpret the latter as consistent with the presence of potential diseconomies of scale for overlapping projects.

5.2 Second step: estimation of complementarities

Let $\kappa^{\omega}_i(Z, W, X)$ denote the complementarity bidder i assigns to outcome ω at observables (Z, W, X). We consider the following simple linear specification for $\kappa_i(\cdot)$:

$$\kappa^{\omega}_i(Z, W, X; \theta_0) = \mathbf{C}^{\omega}(Z_i, W_i) \cdot \theta_0, \tag{10}$$

[20]While the parametrization of Σ_{ilt} does not imply its positive semi-definitiveness, the estimated variance-covariance matrix is positive semi-definite.

Table 5: First-step MLE estimates of parameters in G_i

Mean μ_{ilt}	$\hat{\alpha}$	MLE SEs	95% CI	
Constant	0.3208	0.0159	0.2896	0.352
Log engineer's estimate	0.9809	0.0009	0.9791	0.9827
Log rivals in auction	-0.0422	0.0027	-0.0475	-0.0369
Multiple bids dummy	-0.1333	0.0207	-0.1739	-0.0927
Log sum engineer's (across l)	0.0079	0.0014	0.0052	0.0106
Log sum rivals (across l)	-0.0078	0.0022	-0.0121	-0.0035
Log days to the start	0.0032	0.0008	0.0016	0.0048
Standardize backlog	0.0036	0.001	0.0016	0.0056
Same-type-auctions index	-0.0214	0.005	-0.0312	-0.0116
Fraction overlapping time	0.0186	0.0035	0.0117	0.0255
Log number of big rivals faced	0.0056	0.0022	0.0013	0.0099
Log number of regular rivals faced	0.0242	0.0024	0.0195	0.0289
Big bidder	0.008	0.0045	-0.0008	0.0168
Regular bidder	-0.0056	0.0025	-0.0105	-0.0007
Log distance to project	0.0146	0.0009	0.0128	0.0164
Log distance across played projects	0.0074	0.0017	0.0041	0.0107
Bidder Type FE	YES	-	-	-
Auction Type FE	YES	-	-	-
Year FE	YES	-	-	-
Month FE	YES	-	-	-
Variance σ^2_{ilt}	**$\hat{\beta}$**	**MLE SEs**	**95% CI**	
Constant	0.0984	0.0743	-0.0472	0.244
Multiple bids dummy	-0.2032	0.019	-0.2404	-0.166
Log engineer's estimate	-0.2647	0.0054	-0.2753	-0.2541
Covariance ρ_{iklt}	**$\hat{\gamma}$**	**MLE SEs**	**95% CI**	
Constant	0.2057	0.023	0.1606	0.2508
Same county projects	0.2209	0.0275	0.167	0.2748
Same type projects	0.1261	0.0188	0.0893	0.1629
Fraction overlapping time	-0.0275	0.0203	-0.0673	0.0123

where $\theta_0 \subset \Theta$ is a $p \times 1$ vector of unknown parameters and $\mathbf{C}^{\omega}(Z_i, W_i)$ is a known $p \times 1$ function of (Z_i, W_i) describing characteristics relevant to combination ω.[21]

In view of our linear specification (10) for κ_i, we here consider estimation under the following combination of Assumptions 5 and 8:

$$F_{i,m}(\cdot|Z, W, X) = F(\cdot|Z_i, X_l) \text{ for all } i, \text{ all } m \in \mathcal{M}_{\rangle}, \text{ and all } (Z, W, X), \tag{11}$$

where $F_{i,m}(\cdot|\cdot)$ denotes the conditional marginal distribution of bidder i's standalone valuation for object $m \in \mathcal{M}_i$. We thus exploit both variation in rival characteristics Z_{-i} and variation in combination characteristics W as defined in section 3 above.

The essence of our identification strategy is to compare the equilibrium bidding behavior of *similar bidders* competing for *similar contracts* within letting environments that differ either in rival characteristics Z_{-i} or in combination characteristics W. We implement this intuition as follows.

(a) Given a bidder i, randomly select an auction m played by i.

(b) Holding auction m and the number of auctions played by i constant, draw two different sets of auctions played by bidders of the same type as i. This yields two distinct letting environments j and k, which by construction are identical on non-excluded dimensions but differ on excluded dimensions.

(c) For each of these hypothetical letting environments, we approximate the key equilibrium-dependent terms $\Psi_i(Z, W, X)$ and $\Upsilon_i(Z, W, X)$ via simulation as follows. We first draw a size-R sample of bids $\{b_i^r\}_{r=1}^R$ from the estimated conditional distribution $\hat{G}_i(\cdot|Z, W, X)$ of bids submitted in lettings with the characteristics given.[22] For each draw b_i^r, we then compute estimates of $\Gamma_i(b_i^r|Z, W, X)$, $\nabla\Gamma_i(b_i^r|Z, W, X)$, and $\nabla P_i(b_i^r|Z, W, X)$ based on our initial estimate of $\hat{G}_i(\cdot|Z, W, X)$.

[21] In practice, $\mathbf{C}^{\omega}(\cdot)$ includes the following elements: bidder i's net backlog after winning combination ω (i.e. i's current backlog plus the sum of engineer's estimates among projects won), the Herfindahl index of auction types in ω, the fraction of time overlap among projects in ω, the product of fraction overlapping time and total combination size, the distance among projects in combination ω, and a set of dummies for bidder size (large, regular) and bidder type. Construction of these variables is described in detail in Section 4.1 above.

[22] In practice we set R to 500, with larger samples having very little effect on results.

Taking appropriate averages of these, we ultimately obtain the desired approximations to Ψ and Υ:

$$\hat{\Upsilon}_i(Z,W,X) = \frac{1}{R}\sum_{r=1}^{R} b_i^r + \nabla\Gamma_i(b_i^r|Z,W,X)^{-1}\,\Gamma_i(b_i^r|Z,W,X);$$
$$\hat{\Psi}_i(Z,W,X) = \frac{1}{R}\sum_{r=1}^{R} \nabla\Gamma_i(b_i^r|Z,W,X)^{-1}\nabla P_i(b_i^r|Z,W,X)^T.$$

(d) Select the elements in $\Upsilon_i^j, \Psi_i^j, \Upsilon_i^k, \Psi_i^k$ corresponding to auction m for which the exclusion restrictions implied by Assumption 8 are satisfied.

(e) Repeat the previous steps from (a) to (d)

(f) Construct an estimator $\hat{\theta}$ for θ_0 by minimizing violations of (9) as measured by the following least-squares estimation criterion:

$$\min_{\theta\in\Theta}\sum_i\sum_m \left(\Upsilon_{i,m}^j - \Upsilon_{i,m}^k - \left(\Psi_{i,m}^j\mathbf{C}_{i,m}^j - \Psi_{i,m}^k\mathbf{C}_{i,m}^k\right)\cdot\theta\right)^2. \tag{12}$$

While so far we have emphasized first moments as sufficient for identification, the analysis in Section 3 in fact yields a much stronger identifying restriction: invariance of the *whole distribution* of V_{im} to suitable variation in (Z_{-i}, W). Both in principle and in practice, matching on features beyond simple first moments conveys substantial additional information on the shape of this distribution, thereby significantly improving precision of estimates of θ_0. To incorporate this additional information, we extend the criterion (12) to enforce invariance also in certain predicted quantiles of V_{im}.[23] Since these predicted quantiles are themselves functions of θ, we minimize this richer criterion by iteration. Starting from the initial estimate $\hat{\theta}$ derived from (12), we simulate a new set of predicted means and quantiles of V_{im} as above. We then minimize divergence in these to obtain a new estimate $\hat{\theta}_1$ of θ_0, and iterate this procedure until convergence. In each iteration, the criterion to minimize reduces to

[23] In practice, we consider the 25^{th}, 50^{th} and 75^{th} quantiles.

a linear least-squares estimator (OLS), which we implement via robust regression to reduce influence of skewness, outliers, and non-constant variance in the errors. Standard error are boot-strapped.[24]

5.3 The main result: structural estimates of θ_0

Table 6 reports estimates $\hat{\theta}$ derived from mapping the first-step estimates $\hat{G}_1, ..., \hat{G}_N$ in Section 5.1 through the second-step procedure outlined in Section 5.2. Although we estimate $\hat{G}_1, ..., \hat{G}_N$ for all bidders, when forming the criterion (12) used to estimate $\hat{\theta}$ we restrict attention to the subsample of bidders competing in two auctions. This restriction serves both to reduce computational costs and to improve the numerical quality of our simulated criterion (12); as usual, performance on both fronts declines rapidly in higher dimensions.[25] Note, however, that as defined above $\kappa(\cdot)$ depends only on characteristics such as total size of, overlap between, and distance among projects in a given combination. Insofar as these scale naturally to other combination sizes, so will our estimates of the complementarity vector $\kappa(\cdot)$.

Bearing in mind that positive signs reflect "positive complementarities" (lower costs) while negative signs reflect "negative complementarities" (higher costs), the coefficients reported in Table 6 have the following economic interpretations. The variable "Current backlog plus sum of engineer's estimates" reflects the *ex post backlog* that i would realize in the event of a combination win, with a negative coefficient on this variable suggesting that higher ex post backlog renders a joint win less valuable, as we would expect in the presence of capacity constraints. The variables "Fraction overlapping time" and "Fraction overlapping time $\times$ Sum engineer's estimates" measure the extent to which project schedules overlap, with negative signs on these suggesting that schedule overlap substantially increases both average completion costs and the rate at which completion costs increase in combination size. Meanwhile, signs on the variables "Distance among projects" and "Same-type

[24]In practice, we generate 1297 pseudo-observations as described above to construct our criterion function.

[25]This is particularly true with respect to simulated gradients of probabilities of higher-order combinations, which would play a central role in any attempt to estimate $\kappa(\cdot)$ in higher dimensions.

Table 6: Estimated complementarity parameters θ_0

Combination characteristics (Elements of W)	$\hat{\theta}$	SE
Current backlog in '000 plus sum engineer's estimates in '000	-0.0013	0.0003
Fraction overlapping time among projects in combination	-21.1490	7.0028
Fraction overlapping time $\times$ Sum engineer's estimates in '000	-0.0063	0.0014
Distance in KM among projects in combination	-0.0288	0.0062
Same-type-auctions index	82.8348	17.5152
Bidder type / size FE	YES	-

Units are in thousands of dollars, positive κ means lower cost (more cost synergy, larger complementarities) between projects.

auction index" suggest that both greater distance and greater heterogeneity make projects more substitutable. Finally, although not reported in Table 6, we include a vector of bidder type and bidder size dummies; signs on these vary, but suggest a positive intercept for $\kappa(\cdot)$ on aggregate as we quantify in detail next.

To illustrate the economic significance of these parameter estimates, we next translate the parameter estimates $\hat{\theta}$ in Table 6 into estimates for the underlying complementarities $\kappa(\cdot)$ themselves. These will of course vary both across bidders and across combinations, so for the moment we proceed as follows. We first construct, for each bidder i in the sample, the complementarity $\kappa^{\omega}(Z_i, W_i; \hat{\theta})$ associated with the largest combination played by i. We then normalize this complementarity by the total size of the relevant combination, and analyze the distribution of these normalized complementarities across bidders.

Table 7 summarizes the results of this procedure, reporting quantiles of normalized complementarities for both (i) all bidders competing in two auctions and (ii) all bidders in our MDOT sample. As evident from Table 7, there is substantial heterogeneity in complementarities across bidders in the MDOT sample, with a joint win leading to cost *savings* of approximately 13 percent of combination size at the 95th (best) quantile of complementarities, transitioning to cost *increases* of approximately 3.5 percent at the 5th (worst) quantile. Recalling the parameter estimates in Table

6, we view these pattern as consistent with an underlying U-shaped cost curve, with completion costs falling until firm resources are fully employed and rising thereafter.

We conclude this section with a note on interpretation of Tables 6 and 7 under endogenous entry. In Appendix B, we embed the bidding model considered here within a fully specified entry and bidding game, showing that our estimation strategy is robust to this extension. Hence the parameter estimates reported in Table 6 remain valid even under entry. In interpreting Table 7, however, endogenous entry will be pivotal: the distribution of complementarities among projects *in which bidders enter* will obviously differ from that which would arise if projects were randomly assigned. In particular, insofar as bidders tend to bid for complementary combinations, we would expect the distribution in Table 7 to be positively selected.

6 Counterfactuals

While the simultaneous first-price auction is clearly inefficient when bidders have combinatorial preferences, little is known empirically about the magnitude of these inefficiencies in practice. Furthermore, little is known (either theoretically or empirically) about the revenue properties of the simultaneous first-price auction relative to other feasible multi-object mechanisms such as the Vickery-Clarke-Groves (VCG) mechanism, the combinatorial proxy auction (Ausubel and Milgrom 2002), or the clock-proxy auction (Ausubel, Crampton and Milgrom 2006). Given that implementation of such combinatorial mechanisms involves substantial practical costs (even solving the allocation problem once is NP-hard), determining the magnitude of their potential revenue and efficiency effects is crucial in evaluating whether policymakers might want to switch. If efficiency gains are small and / or revenue effects are ambiguous, an optimal policymaker may prefer the simplicity and transparency of the simultaneous first-price auction to better-performant but more complex combinatorial mechanisms. Conversely, if large efficiency and / or revenue gains are feasible, incurring greater combinatorial implementation costs may be worthwhile.

In this section, we compare revenue and efficiency outcomes of the simultaneous low-bid first-price auction with those of a descending combinatorial proxy auction

Table 7: Distribution of complementarities across bidders

	Quantile of normalized $\hat{\kappa}^{\omega}(Z_i, W_i)$ in:	
Quantile rank	*Two-auction sample*	*Full sample*
5th	-0.0352	-0.0365
10th	-0.0205	-0.0239
15th	-0.0103	-0.0170
20th	-0.0045	-0.0117
25th	-0.0010	-0.0079
30th	0.0022	-0.0050
35th	0.0057	-0.0023
40th	0.0094	0.0007
45th	0.0139	0.0034
50th	0.0187	0.0062
55th	0.0231	0.0097
60th	0.0287	0.0139
65th	0.0358	0.0186
70th	0.0437	0.0242
75th	0.0554	0.0321
80th	0.0732	0.0414
85th	0.0923	0.0554
90th	0.1244	0.0812
95th	0.1847	0.1291

"Normalized $\hat{\kappa}^{\omega}(Z_i, W_i; \hat{\theta})$" denotes estimated complementarity $\hat{\kappa}^{\omega}(Z_i, W_i; \hat{\theta})$ among projects bid by i divided by the sum of engineer's estimates among projects bid by i, with quantiles evaluated over the empirical distribution of (Z_i, W_i) over all bidders and periods in the sample indicated. Positive fractions represent positive complementarities (lower costs). Thus the statement that the 50th quantile of normalized $\hat{\kappa}^{\omega}(Z_i, W_i)$ is 0.0075 in the full sample means that for the median bidder a joint win would generate cost savings equal to approximately 0.7 percent of combination size, and similarly.

a la Ausubel and Milgrom (2002). While efficiency can also be achieved with the VCG mechanism, this can also exhibit very poor revenue performance. The Ausbel-Milgrom proxy auction mitigates the potential revenue disadvantages of the VCG auction, while still achieving efficiency so long as bidders report their true preferences to the proxy agent.

Descending proxy auction Adapted to our procurement setting, the descending proxy auction operates as follows. First, each bidder i reports to its proxy agent a $(2^{M_i} - 1) \times 1$ vector describing costs of completion for each possible combination of the M_i products on which i has undertaken cost discovery. Second, proxies compete on behalf of bidders in a virtual descending package auction, bidding according to the following rule: in each bidding round, submit the allowable package bid that, if accepted, would maximize the bidder's profit given its reported costs. After each bidding round, a provisional winning allocation is determined by minimizing procurer costs over existing bids, and bidding proceeds to the next round. If no new bids are submitted in a round, the auction ends.

Consistent with most prior work on proxy auctions, we restrict attention to the case where bidders truthfully report costs. This guarantees that the final allocation is efficient and in the core of the corresponding exchange game. Note, however, that it is uncertain whether truthful reporting is an equilibrium in general.[26] Insofar as false reports distort final allocations, our results may overstate gains from the proxy auction. Nevertheless, we see truthful revelation as a useful and practical benchmark for comparison with the simultaneous first-price auction.

Computation of final outcomes in the Ausubel-Milgrom proxy auction is known to be extremely challenging, requiring one to solve a NP-hard winner determination problem for every bidding round. Since the proxy auction obtains (approximate) efficiency only with a small bid increment, and the number of bidding rounds required for convergence increases substantially as the bid increment decreases, naive application of the Ausubel-Milgrom algorithm can be extremely costly computationally. We therefore focus instead on two variants of the Ausubel-Milgrom auction iden-

[26]See related discussion in, e.g., Ausubel and Milgrom (2002).

tified by Sandholm (2006) as having good computational properties: the *safe-start proxy auction*, in which starting bids for each bidder are determined by the VCG payment rule, and the *increment scaling proxy auction*, in which the bid increment automatically scales down as the auction proceeds. In both variants we target a final-iteration bid increment of $1000, which is quite small as bids are typically in hundreds of thousands to millions of dollars. These algorithms need not generate the same revenue as the naive proxy auction, but retain its desirable efficiency and revenue properties. See Sandholm (2006) for detailed discussion of these algorithms.

Counterfactual implementation As described above, the main challenge in implementing our counterfactuals is computational: the optimal winner determination problem in combinatorial auctions is well-known to be NP-hard, with complexity growing very rapidly in the number of auctions and bidders.[27] To ease this computational burden, we restrict our counterfactual sample as follows. First, we drop the 5 percent of bidders submitting more than 8 bids. Second, starting from the 778 self-contained lettings in our counterfactual sample, we partition each letting involving more than one million possible allocations into smaller sub-lettings via the Girvan-Newman algorithm: interpreting each letting as a network with bidders and auctions defining nodes and bids defining edges, we iteratively drop bids with the highest "edge connectivity" until the letting is partitioned.[28] We then repeat this process until no letting involves more than 1 million potential allocations. Our final counterfactual sample thus involves 1656 self-contained lettings representing roughly 95 percent of unique bidders and roughly 70 percent of total bids, of which 1193 lettings (our primary interest) involve at least two auctions.

Given this sample, we implement our counterfactual comparisons as follows.

[27]For instance, to determine the efficient allocation in a letting with 30 auctions receiving 4 bids each – by no means the largest in our sample – we might have to compare up to $4^{30} \approx 10^{18}$ possible allocations. Even comparing a billion allocations per second – faster than feasible on a standard microprocessor – this would take more than 30 years to solve exactly. Real-world approaches to solving large-scale combinatorial auctions rely instead on heuristic winner-determination algorithms, which aim to achieve reasonable solution quality in reasonable computation time. Since we here wish to solve exactly, we do not explore these.

[28]See Girvan and Newman (2002) for a formal description of this algorithm.

First, for each bidder i and letting t in the counterfactual sample, we draw a sample of bids $\{b_i^{rt}\}_{r=1}^{R}$ from the corresponding bid distribution $\hat{G}_i(\cdot)$ estimated in Step 1 of our structural procedure. Second, for each bid vector b_i^{rt} drawn for each bidder i, we recover the corresponding standalone valuation vector v_i^{rt} implied by the inverse bid function (2), taking as given the estimates $\hat{\kappa}_i(\cdot)$ for $\kappa_i(\cdot)$ obtained in Step 2 of our structural procedure. For each letting t in the counterfactual sample and each replication $r \in \{1, ..., R\}$, we then proceed in three steps.

First, we simulate the allocation a_{FPA}^{rt} and procurement cost C_{FPA}^{rt} arising under the simultaneous first-price auction given bid realizations $\{b_i^{rt}\}_{i=1}^{N}$; i.e. awarding each auction to the bidder submitting the lowest standalone bid. Then, given estimated complementarities $\{\hat{\kappa}_i(\cdot)\}_{i=1}^{N}$ and estimated valuations $\{v_i^{rt}\}_{i=1}^{N}$, we simulate total social costs of project completion S_{FPA}^{rt} corresponding to allocation a_{FPA}^{rt}.

Second, assuming truthful reporting of types by bidders, we simulate proxy auction procurement costs C_{PROXY}^{rt} based on the safe-start and incremental scaling algorithms described above. In both variants, we target a final iteration bid increment of \$1000, which is quite small relative to typical bids. While in principle efficiency in proxy auctions obtains only when the bid increment approaches zero, in practice our \$1000 bid increment leads to allocations which are virtually indistinguishable from the VCG mechanism.

Revenue and efficiency We now describe results of this counterfactual comparison based on $R = 10$ simulation replications, focusing on the subsample of lettings involving at least two auctions. For purposes of these simulations, we set MDOT's effective reserve price for each project equal to 125 percent of the MDOT engineer's cost estimate; other plausible values generate very similar results.

Two striking patterns emerge from this exercise. First, as expected, the simultaneous first-price mechanism is socially inefficient, generating expected social costs of roughly \$6.60 million per counterfactual letting versus \$6.36 million per letting for the (socially efficient) proxy mechanism. In level terms this difference is nontrivial, translating to an average savings of roughly \$40,000 per auction. Yet in percentage terms gains are relatively small: roughly 3.7 percent social cost savings relative to total project completion costs under the simultaneous first-price mechanism.

Second, and even more striking, expected payments by MDOT to bidders are very similar across mechanisms, with the proxy auction generating savings of only about one percent of total MDOT payments. We emphasize that this is not a prediction of the theory; with different parameters, one can easily obtain substantial differences in revenue.

Recall that (by construction) the number of bids submitted per bidder is lower in our counterfactual than in the data: approximately 2.2 bids per bidder in the counterfactual, versus 2.7 bids per bidder in the data. To determine whether our results are sensitive to this, we reweight counterfactual lettings such that the average number of bids per bidder in the counterfactual equals the average number of bids per bidder in the data.[29] This increases estimated efficiency gains slightly to approximately 4.3 percent, leaving estimated MDOT payments essentially unchanged.

Finally, to explore the role of complementarities per se in the performance of the simultaneous first-price mechanism, we compute a measure of "complementarity importance" defined as follows: for each letting, we find the maximum (in absolute value) complementarity among combinations played by each bidder, average this measure across bidders, and then normalize by average size of projects. Reweighting the bid-adjusted sample by this measure of complementarity importance, estimated gains increase to about 5.5 percent social savings and 1.5 percent MDOT savings relative to the corresponding simultaneous first-price baseline. In other words, counterfactual gains tend to be substantially larger in lettings where complementarities are more important – as expected given the discussion above.

On the whole, we view these findings as strong suggestive evidence that the simultaneous first-price mechanism in fact performs remarkably well in the MDOT marketplace. This analysis is of course only partial in that we effectively hold entry behavior fixed across mechanisms. By construction, social savings not captured by MDOT must accrue as profit to bidders, and in equilibrium this should translate

[29]Specifically, we choose a final weight vector to minimize the Euclidean distance between our final weights and unit weighting, subject to the constraint that the average number of bids per bidder in the reweighted sample equals that in the data.

into greater entry. This in turn might generate slightly larger revenue effects than we find here. In contrast, since new entrants are by definition marginal, we expect true efficiency gains to be similar to those reported above.

7 Conclusion

Motivated by an institutional framework common in procurement applications, we develop and estimate a structural model of bidding in simultaneous first-price auctions, to our knowledge the first such in the literature. Non-parametric and semi-parametric identification of the model is achieved under standard exclusion restrictions. Finally, we apply this framework to data on Michigan Department of Transportation highway construction and maintenance auctions. Our estimates suggest that winning a two-auction combination generates cost effects ranging from roughly 3.5 percent cost increases (relative to combination size) at the 5th percentile to roughly 13 percent cost savings (relative to combination size) at the 95th percentile, with combination costs increasing in joint size of, scheduling overlap between, and distance between projects in the combination. Building on these observations, we compare performance of the simultaneous first-price mechanism with performance of a descending proxy auction a la Ausubel and Milgrom (2002). Despite the presence of substantial complementarities (both positive and negative) in the data, we find that this alternative mechanism generates relatively modest gains: roughly four percent savings in social costs of project completion, with little change in MDOT's expected procurement costs. We view this as strong suggestive evidence that simultaneous first-price auctions can perform relatively well even in environments with economically important complementarities. This observation may partially rationalize the widespread popularity of simultaneous first-price auctions in practice.

References

ATHEY, S., J. LEVIN, and E. SEIRA (2011), "Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions." *The Quarterly Journal of Economics*,

126, 207–257.

AUSUBEL, L.M., and P. CRAMTON (2004), "Auctioning Many Divisible Goods", *Journal of the European Economic Association* **2**, 480–493, April-May.

AUSUBEL, L.M., P. CRAMTON, P. MCAFEE, and J. MCMILLAN (1997), "Synergies in Wireless Telephony: Evidence from the Broadband PCS Auctions", *Journal of Economics and Management Strategy* **6(3)**, 497–527.

AUSUBEL, L.M., P. CRAMTON, and P. MILGROM (2006), "The Clock-Proxy Auction: A Practical Combinatorial Auction Design", in P. Cramton, Y. Shoham, and R. Steinberg (eds.), *Combinatorial Auctions*, MIT Press, MIT Press, Chapter 5, 115–138.

BAJARI, P., and J. FOX (2013), "Measuring the Efficiency of an FCC Spectrum Auction", *American Economic Journal: Microeconomics* **5(1)**, 100–146.

BALAT, J. (2015), "Highway Procurement and the Stimulus Package: Identification and Estimation of Dynamic Auctions with Unobserved Heterogeneity." Working paper, Johns Hopkins University.

CANTILLON, E., and M. PESENDORFER (2006), "Combination Bidding in Multi-Unit Auctions." Working paper, London School of Economics.

CHAPMAN, J.T., D. MCADAMS, and H.J. PAARSCH (2007), "Bounding Revenue Comparisons across Multi-Unit Auction Formats under -Best Response." *American Economic Review* **97(2)**, 455–458.

CRAMTON, P. (2006), "Simultaneous Ascending Auctions." Ch. 4 in Cramton, P., Shoham, Y., and Steinberg, R. (eds) *Combinatorial Auctions* (Cambridge, MA: MIT Press), 99–114.

CRAMTON, P., Y. SHOHAM, and R. STEINBERG (eds) (2006), *Combinatorial Auctions* (Cambridge, MA: MIT Press).

DE SILVA, D. (2005), "Synergies in Recurring Procurement Auctions: An Empirical Investigation." *Economic Inquiry* **43(1)**, 55–66.

DE SILVA, D., T. JEITSCHKO, and G. KOSMOPOULOU (2005), "Stochastic Synergies in Sequential Auctions." *International Journal of Industrial Organization* **23**, 183-201.

DE VREIS, S., and R.V. VORHA (2003), "Combinatorial Auctions: A Survey." *INFORMS Journal of Computing* **13(3)**, 284–309.

FELDMAN, M., H. FU, N. GRAVIN, and B. LUCIER, (2012), "Simultaneous Auctions are (Almost) Efficient." (Working Paper, *http://arxiv.org/abs/1209.4703*).

FLAMBARD, V., and I. PERRIGNE (2006), "Asymmetry in Procurement Auctions: Evidence from Snow Removal Contracts." *The Economic Journal* **(116), Issue 514**, 1014–1036.

FEVRIER, P., R. PREGET, and M. VISSER (2004), "Econometrics of Share Auctions." *Center for Research in Economics and Statistics (CREST) Working Paper*, 2002–2009.

GENTRY, M., T. KOMAROVA, P. SCHIRALDI, and W. SHIN (2016), "On Monotone Equilibria in Simultaneous First-Price Auctions." Working paper, London School of Economics.

GIRVAN, M. and M. E. NEWMAN (2002), "Community structure in social and biological networks." Proceedings of the National Academy of Sciences, 99, 78217826.

GHOSH, G. (2012), "Simultaneous First-Price Auctions with Budget-Constrained Bidders." Working paper.

GROEGER, J.R. (2014), "A Study of Participation in Dynamic Auctions." *International Economic Review* 55(4), 1129-1154.

GUERRE, E., I. PERRIGNE, and Q. VUONG (2000), "Optimal Nonparametric Estimation of First-Price Auctions." *Econometrica*, **68(3)**, 525–574.

HENDRICKS, K., J. PINKSE, and R.H. PORTER (2003), "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions." *Review of Economic Studies*, **70**, 115–145.

HENDRICKS, K., and R.H. PORTER (1988), "An Empirical Study of an Auction with Asymmetric Information." *American Economic Review*, **78(5)**, 865–883.

HORTACSU, A. (2011), "Recent Progress in the Empirical Analysis of Multi-Unit Auctions." *International Journal of Industrial Organization*, **29(3)**, 345–349.

HORTACSU, A. and D. McADAMS (2010), "Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market." *Journal of Political Economy* , **188(5)**, 833–865.

HORTACSU, A., and S.L. PULLER (2008), "Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market." *RAND Journal of Economics*, **39(1)**, 86–114.

JACKSON, M.O., L. K. SIMON, J.M. SWINKELS, and W.R. ZAME (2002), "Communication and Equilibrium in Discontinuous Games of Incomplete Information." *Econometrica*, **70(5)**, 1711–1740.

JOFRE-BONET, M., and M. PESENDORFER (2003), "Estimation of a Dynamic Auction Game." *Econometrica*, **71(5)**, 1443–1489.

KAWAI, K. (2011), "Auction Design and the Incentives to Invest: Evidence from Procurement Auctions." Working paper, UCLA.

KIM, S.W., M. OLIVARES and G.Y. WEINTRAUB (2014), "Measuring the Performance of Large-Scale Combinatorial Auctions: A Structural Estimation Approach." *Management Science* **60(5)**, 1180-1201.

KLEMPERER, P. (2008), "A New Auction for Substitutes: Central Bank Liquidity Auctions, the U.S. TARP, and Variable Product-Mix Auctions. Bayesian Games and the Smoothness Framework." Working paper, Oxford University.

KLEMPERER, P. (2010), "The Product-Mix Auction: A new auction design for differentiated goods." *Journal of the European Economic Association*, **8(2-3)**, 526–536.

KRASNOKUTSKAYA, E. (2011), "Identification and Estimation in Procurement

Auctions under Unobserved Auction Heterogeneity." *Review of Economic Studies* **78(1)**, 293–327.

KRASNOKUTSKAYA, E., and K. SEIM (2011), "Bid Preference Programs and Participation in Procurement." *American Economic Review* **101(6)**, 2653–2686.

KRISHNA, V., and R. W. ROSENTHAL (1996), "Simultaneous Auctions with Synergies." *Games and Economic Behavior* **17**, 1–31.

LI, T., and B. ZHANG (2010), "Testing For Affiliation In First-Price Auctions Using Entry Behavior." *International Economic Review* **51(3)**, 837–850.

LI, T., and B. ZHANG (2015), "Affliation and Entry in First-Price Auctions with Heterogeneous Bidders: An Analysis of Merger Effects." *American Economic Journal: Microeconomics* **7(2)**, 188–214.

LI, T. and X. ZHENG (2009), "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions." *Review of Economic Studies* **76(4)**, 1397–1429.

LI, T. and X. ZHENG (2012), "Information Acquisition and/or Bid Preparation: A Structural Analysis of Entry and Bidding in Timber Sale Auctions." *Journal of Econometrics* **168(1)**, 29–46.

LU, J. and PERRIGNE, I. (2008), "Estimating Risk Aversion Using Ascending and Sealed Bid Auctions: Application to Timber Data." *Journal of Applied Econometrics* **23**, 871–896.

LUNANDER, A., and S. LUNDBERG (2013), "Bids and Costs in Combinatorial and Noncombinatorial Procurement Auctions – Evidence from Procurement of Public Cleaning Contracts." *Contemporary Economic Policy* **31(4)**, 733–745.

MILGROM, P. (2000a), "Putting Auction Theory to Work: The Simultaneous Ascending Auction." *Journal of Political Economy* **108(2)**, 245–272.

MILGROM, P. (2000b), "Putting Auction Theory to Work: Ascending Auctions with Package Bidding." Working paper, Stanford University.

MILGROM, P., and R. WEBER (1985), "Distributional Strategies for Games with Incomplete Information." *Mathematics of Operations Research* **10(4)**, 619–632.

MORETON, P. and P. SPILLER (1998), "Whats in the Air: Interlicense Synergies and their Impact on FCC Broadband PCS Spectrum Auctions." *Journal of Law and Economics* **41(2)**, 677–716.

Ofcom (2007). Digital dividend review: A statement on our approach to awarding the digital dividend. United Kingdom Office of Communications. http://www.ofcom.org.uk/consult/condo statement/

REGUANT, M. (2014), "Complementary Bidding Mechanisms and Startup Costs in Electricity Markets", *Review of Economic Studies* **81(4)**, 1708–1742.

ROSENTHAL, R.W., and WANG, R. (1996), "Simultaneous Auctions with Synergies and Common Values." *Games and Economic Behavior* **17**, 32–55.

SANDHOLM, T. (2006), "Optimal Winner Determination Algorithms." Chapter 14 in *Combinatorial Auctions*, CRAMTON, SHOHAM, and STEINBERG, R. (eds), Cambridge, MA: MIT Press.

SOMAINI, P. (2013), "Competition and Interdependent Costs in Highway Procurement." Working Paper, Stanford University.

SYRGKANIS, V. (2012), "Bayesian Games and the Smoothness Framework", (Working Paper, *http://arxiv.org/abs/1203.5155*).

SZENTES, B., and ROSENTHAL, R.W. (2003), "Three-object Two-bidder Simultaneous Auctions: Chopsticks and Tetrahedra." *Games and Economic Behavior* **44**, 114–133.

WOLAK, F. (2007), "Quantifying the Supply-Side Benefits from Forward Contracting in Wholesale Electricity Markets." *Journal of Applied Econometrics* **22(7)**, 1179–1209.

Appendix A: Proofs

Proof of Proposition 1. The proof of Proposition 1 rests on two key claims. First, the first-order system (2) must be well-defined for almost every b_i submitted by i, i.e. almost everywhere with respect to the measure induced by $G_i(\cdot|Z, W)$. Second, at almost every b_i at which first order conditions hold, the first-order system (2) must be invertible. We establish each claim in turn.

First show that the first order system (2) is well-defined for almost every b_i submitted by i. Recall that we can write bidder i's objective as

$$\pi(v_i, b|K; Z, W) = (\Omega v_i + K - \Omega b)^T P_{-i}(b|Z, W).$$

where v_i and K are given at the time of maximization. Note that the system (2) necessarily holds at any best respose where $\pi(v_i, \cdot|K; Z, W)$ is differentiable and that Assumption 2 implies that each observed b_i is a best response. Hence the system (2) will be well defined for almost every b_i submitted by i if and only if $\pi(v_i, \cdot|K; Z, W)$ is differentiable almost everywhere with respect to the measure on B_i induced by $G_i(\cdot|Z, W)$. But under Assumption 3, $G_i(\cdot|Z, W)$ is absolutely continuous. To establish the claim, it thus suffices to show differentiability of $\pi(v_i, \cdot|K; Z, W)$ a.e. with respect to Lebesgue measure on $\mathcal{B}_i$.

Clearly $(\Omega v_i + K - \Omega b)$ is differentiable in b. Thus differentiability of $\pi(v_i, \cdot|K; Z, W)$ at b is equivalent to differentiability of $P_{-i}(\cdot|Z, W)$ at b. Let B_{-i} be the $M_i \times 1$ random vector describing maximum rival bids in the set of auctions in which i participates. Again applying Assumption 3 to rule out ties, the probability i wins combination ω at bid b is

$$P^{\omega}(b|Z, W) = \Pr(\{\cap_{\{m:\omega_m=1\}} 0 \leq B_{-i,m} \leq b_{i,m}\} \cap \{\cap_{\{m:\omega_m=0\}} b_{i,m} \leq B_{-i,m} < \infty\}).$$

For each $\omega \in \Omega_i$, let b^{ω} be the $(\sum \omega) \times 1$ sub-vector of b describing i's bids for objects in ω, B^{ω}_{-i} be the $(\sum \omega) \times 1$ sub-vector of B_{-i} describing maximum rival bids for objects in ω, and $G^{\omega}_{-i}(b^{\omega}|Z, W)$ be the equilibrium joint c.d.f. of B^{ω}_{-i} at (Z, W). Applying the formula for a rectangular probability and simplifying, we can then represent $P_{-i}(\cdot|Z, W)$ in the form

$$P^{\omega}_{-i}(b|Z, W) = \sum_{\omega' \in \Omega} a^{\omega}_{\omega'} G^{\omega'}_{-i}(b^{\omega'}|Z, W),$$

where each $a^{\omega}_{\omega'}$ is a known scalar (determined by ω, ω') taking values in $\{-1, 0, 1\}$. But by absolute continuity each c.d.f. $G^{\omega}_{-i}(\cdot|Z, W)$ is differentiable a.e. (Lebesgue) in its support, and interpreted as a function from $\mathcal{B}_i$ to $\mathbb{R}^{M_i}$, each $b^{\omega'}$ is continuously differentiable in b. Thus interpreted as a function from $\mathcal{B}_i$ to $\mathbb{R}$, each $G^{\omega'}_{-i}(b^{\omega'}|Z, W)$ is differentiable on a set of full Lebesgue measure in B_{-i}. The set of points in $\mathcal{B}_i$ at which all $G^{\omega'}_{-i}(b^{\omega'}|Z, W)$ are differentiable is the intersection of points in $\mathcal{B}_i$ at which each $G^{\omega'}_{-i}(b^{\omega'}|Z, W)$ is differentiable, i.e. the intersection of a finite collection of sets of full Lebesgue measure in $\mathcal{B}_i$. But from above differentiability of $G^{\omega'}_{-i}(b|Z,W)$ for all ω' implies differentiability of $P^{\omega}_{-i}(b|Z,W)$. Hence $P^{\omega}_{-i}(\cdot|Z,W)$ is differentiable on a set of full Lebesgue measure in $\mathcal{B}_i$. This in turn implies differentiability of $\pi(v_i,\cdot|K;Z,W)$ a.e. with respect to the measure on $\mathcal{B}_i$ induced by $G_i(\cdot|Z,W)$, as was to be shown.

We next establish that the first-order system (2) must yield a unique solution $\tilde{v}$ for almost every b_i submitted by i. Let $\tilde{B}_i$ be the set of points in $\mathcal{B}_i$ at which $\pi(\cdot,\cdot|K;W,Z)$ is differentiable in b; from above, $\tilde{B}_i$ is a subset of full Lebesgue measure in $\mathcal{B}_i$. Choosing any $b \in \tilde{B}_i$ and rearranging (2) yields

$$\nabla_b \Gamma_{-i}(b|Z,W)\tilde{v} = \nabla_b \Gamma_{-i}(b|Z,W)b + \Gamma_{-i}(b|Z,W) - \nabla_b P_{-i}(b|W,Z)^T K_i.$$

Hence uniqueness of $\tilde{v}$ is equivalent to invertibility of the $M_i \times M_i$ matrix $\nabla_b \Gamma_{-i}(b|Z,W)$. Recall that $\Gamma_{-i}(b|Z,W)$ is an $M_i \times 1$ vector whose lth element describes the probability that bid vector b wins auction l. Note that $b \in \tilde{B}_i$ rules out ties at b. Thus for $b \in \tilde{B}_i$ the mth element of $\Gamma_{-i}(b|Z,W)$ is the marginal c.d.f. $G^l_{-i}(b|Z,W)$ of $B_{-i,m}$, from which it follows that $\nabla_b \Gamma_{-i}(b|Z)$ is a diagonal matrix whose m,mth element is the marginal p.d.f. $g_{-i,m}(b|Z,W)$ of $B_{-i,m}$. Hence $\nabla_b \Gamma_{-i}(b|Z,W)$ will be invertible at b if and only if $g_{-i,m}(b|Z,W) > 0$ for all $m = 1,...,M_i$.

But by hypothesis each submitted bid b_i is a best response to rival play at (Z,W) for some (v,K). Suppose that there exists an $\epsilon > 0$ such that $g_{-i,m}(\cdot|Z,W) = 0$ on $(b_{im} - \epsilon, b_i]$. Then player i could infinitesimally reduce b_{im} without affecting either Γ_{-i} or P_{-i}, a profitable deviation for any (v,K). Hence we must have $g_{-i,m}(\cdot|Z,W) > 0$ almost everywhere (Lebesgue) in the support of B_i. By Assumption 3, this in turn implies $g_{-i,m}(\cdot|Z,W) > 0$ for almost every b_i submitted by i. Since m was arbitrary, we must have $\nabla_b \Gamma_{-i}(b_i|Z,W)$ invertible for almost every bid b_i submitted by i. Hence for almost every b_i submitted by i there will exist a unique $\tilde{v}$ satisfying (2) at b_i, given by

$$\begin{aligned}\tilde{v} = b_i &+ \nabla_b \Gamma_{-i}(b_i|Z,W)^{-1}\Gamma_{-i}(b_i|Z,W) \\ &+ \nabla_b \Gamma_{-i}(b_i|Z,W)^{-1}\nabla_b P_{-i}(b_i|W,Z)^T K.\end{aligned}$$

The RHS of this expression is identified up to K, establishing the claim.

□

Appendix B: Entry

In this Appendix, we formally embed the bidding model we describe above within a two-stage entry-plus-bidding model paralleling those considered by Li and Zhang (2015) and Groeger (2014) among others. This discovery process proceeds as follows.

At the beginning of the game, each bidder i is endowed with a $2^L \times 1$ combinatorial valuation vector Y_i drawn by nature from $F_{Y,i}$. However, realizations of Y_i are *ex ante* *unknown* to i and can be discovered by i only through costly entry. Specifically, at the beginning of Stage 1, each bidder i observes a $2^L \times 1$ vector of private *combinatorial entry costs* C_i, with element C_i^S of C_i describing the total cost i must incur to enter auctions $S \in \mathcal{S}$. This cost vector C_i satisfies the following properties:

Assumption 9 (Private Entry Costs)**.** *For each bidder i, C_i is drawn independently of combinatorial preferences Y_i from cost distribution $F_{C,i}$ with support on a compact, convex set $\mathcal{C}_i \subset R^{2^L}$, with C_i private information, $F_{C,i}$ common knowledge, and cost draws independent across bidders: $C_i \perp C_j$ for all i, j.*

Having observed C_i, bidder i chooses a set of auctions $\mathcal{M}_i \in \mathcal{S}$ in which to enter, pays the corresponding entry cost $C_i^{\mathcal{M}_i}$, and proceeds to Stage 2. Then, at the beginning of Stage 2, Bidder i observes the realizations of her combinatorial valuations $Y_i^{S'}$ for all combinations feasible at $\mathcal{M}_i$; that is, for each $S' \in \mathcal{S}$ such that $S' \subset \mathcal{M}_i$. Lastly, bidder i submits a single bid b_{im} for each object m in her entry set $\mathcal{M}_i$. Conditional on realization of any participation structure $\mathcal{M} = \{\mathcal{M}_1, ..., \mathcal{M}_N\}$ realized in Stage 1, the bidding subgame then proceeds exactly as described in the main text.

Following Milgrom and Weber (1985), define a *distributional entry strategy* for player i as a measure ξ_i over $\mathcal{C}_i \times \mathcal{S}$ whose marginal over $\mathcal{C}_i$ is $F_{C,i}$, with $\xi = (\xi_1, ..., \xi_N)$ a profile of distributional entry strategies. Then assuming that at least one Bayes-Nash equilibrium exists, any such equilibrium must have the following form. For each participation structure $\mathcal{M}$ arising from the entry game, let $\Pi(\mathcal{M}) = (\Pi_1(\mathcal{M}), ..., \Pi_N(\mathcal{M}))$ be any vector of candidate bid-stage payoffs in the corresponding bidding subgame. Taking these payoffs as given, let

$$\Xi(S, \xi_{-i}) = E[\Pi_i(S, \mathcal{M}_{-i})|\xi_{-i}]$$

be i's expected net profit from entering auction combination $S \in \mathcal{S}$ given rival entry strategies ξ_{-i} (where the expectation is taken over rival entry sets $\mathcal{M}_{-i}$). We can then write bidder i's Stage 1 problem as:

$$\mathcal{M}_i = \arg\max_{S \in \mathcal{S}} \Xi(S, \xi_{-i}) - C_i^S.$$

The Stage 1 action set for each bidder is the finite set $\mathcal{S}$, and bidders' private entry costs are independent. Hence by Proposition 1 of Milgrom and Weber (1985), there exists an equilibrium in distributional strategies for the entry game corresponding to continuation payoffs $\Pi(\mathcal{M})$. So long as bid-stage payoffs $\Pi(\mathcal{M})$ are themselves generated from play of a Bayes-Nash equilibrium in every bidding subgame, this in turn will constitute an equilibrium of the overall entry and bidding game.

If to this we add the restriction that $F_{C,i}$ is atomless on $\mathcal{C}_i$ for each i, then Proposition 4 of Milgrom and Weber (1985) implies existence of a equilibrium in which bidders play pure entry strategies. Specifically, the set of cost vectors C_i at which bidder i chooses to enter set $S \in \mathcal{S}$, denoted $\mathcal{C}_i^S$, will be the affine cone

$$\mathcal{C}_i^S = \{C_i \in \mathbb{R}^{2^L} : C_i^S - C_i^{S'} \leq \Xi(S, \xi_{-i}) - \Xi(S', \xi_{-i}) \, \forall S' \in \mathcal{S}\},$$

Furthermore, since $C_i \perp Y_i$, equilibrium behavior will involve *variation in participation* by bidder i which is *effectively exogenous* and hence excludable from the perspective of rival bidders. This is precisely the form of variation we exploit in our identification argument.

Appendix C: Partial identification with general G_i

The point identification result for the vector-function of complementarities $\kappa_i(Z_i, W, X)$ and the conditional distribution of $V_i|Z_i, W, X$ relied on the first order conditions obtained from bidder's optimization of the payoff function. To derive those equations we employed the absolute continuity of the bid distribution functions G_i. That, in particular, eliminated the possibility of bidders playing atoms in the equilibrium. In this appendix, we want to illustrate an approach to the identification question without any continuity restrictions imposed on G_i. Our identification method is based on using inequalities for bidder's best responses and employing the exclusion restrictions in Assumption 5 to obtain bounds on $\kappa_i(Z_i, W, X)$ and the CDFs of $V_{il}|Z_i, W, X$. Hereafter we assume that ties are broken independently across auctions.

Let us fix $(Z_i, W, X) \in \mathcal{Z}_i \times \mathcal{W} \times \mathcal{X}$. For each $Z_{-i} \in \mathcal{Z}_{-i}|Z_i, W, X$, bidder maximizes the payoff function

$$\pi(v_i, b; Z, W, X) = v_i^T \Gamma_{-i}(b|Z, W, X) - b^T \Gamma_{-i}(b|Z, W, X) + P_{-i}(b|Z, W, X)^T \kappa_i(Z_i, W, X)$$

with respect to $b \in \mathcal{B}_i$. That is, for each $Z_{-i} \in \mathcal{Z}_{-i}|Z_i, X, W$, every bidder i's bid vector b_i observed in the equilibrium satisfies the inequality

$$\begin{aligned} v_i^T \Gamma_{-i}(b_i|Z_{-i}) - b_i^T \Gamma_{-i}(b_i|Z_{-i}) + P_{-i}(b_i|Z_{-i})^T \kappa(Z_i, W, X) \geq \\ v_i^T \Gamma_{-i}(b|Z_{-i}) - b^T \Gamma_{-i}(b|Z_{-i}) + P_{-i}(b|Z_{-i})^T \kappa(Z_i, W, X) \\ \forall\, b \in \mathcal{B}_i, \quad (13) \end{aligned}$$

where for notational simplicity we wrote $\Gamma_{-i}(\cdot|Z_{-i})$ and $P_{-i}(b_i|Z_{-i})$ instead of $\Gamma_{-i}(b_i|Z, W, X)$ and $P_{-i}(b_i|Z, W, X)$ respectively, thus omitting fixed (Z_i, W, X) from the notation.

Equivalently, (13) can be written as

$$\begin{aligned} v_i^T \left(\Gamma_{-i}(b_i|Z_{-i}) - \Gamma_{-i}(b|Z_{-i})\right) + \left(P_{-i}(b_i|Z_{-i}) - P_{-i}(b|Z_{-i})\right)^T \kappa(Z_i, W, X) \geq \\ b_i^T \Gamma_{-i}(b_i|Z_{-i}) - b^T \Gamma_{-i}(b|Z_{-i}) \quad \forall\, b \in \mathcal{B}_i. \quad (14) \end{aligned}$$

The set of $(v_i, \kappa(Z_i, W, X))$ that satisfy linear inequalities in (14) is clearly convex. If the bidding set $\mathcal{B}_i$ is a continuum, then (14) represents a continuum of linear inequalities in $(v_i, \kappa(Z_i, W, X))$. If the convex set described by (14) is a singleton, then we are in the situation of point identification. Otherwise, we are in a scenario of partial identification. This convex set is fairly difficult to describe in a closed form. A much easier task is to describe its superset and then use it to derive bounds on the c.d.f.s of $V_{il}|Z_i, W, X$.

We start by obtaining a closed form for a superset of the identified set for $\kappa(Z_i, W, X)$. To construct this superset, we consider in (14) only those b that are different from the observed equilibrium vector b_i in one component. Namely, we first consider

$$b = b_i + \epsilon e_m \in \mathcal{B}_i$$

for $\epsilon > 0$ and obtain from (14) that

$$v_{im} \left(\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im} + \epsilon|Z_{-i})\right) + \left(P_{-i}(b_i|Z_{-i}) - P_{-i}(b_i + \epsilon e_m|Z_{-i})\right)^T \kappa(Z_i, W, X) \geq b_{im}\Gamma_{-i}(b_{im}|Z_{-i}) - (b_{im} + \epsilon)\Gamma_{-i}(b_{im} + \epsilon|Z_{-i}),$$

where we used the assumption that the ties are broken independently across auctions at b_i, and thus, a change in the mth component of b_i affects only the mth component of Γ_{-i}. Noting that $\Gamma_{-i,m}$ is (weakly) increasing in b_{im}, we have $\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im} + \epsilon|Z_{-i}) \leq 0$. If bidder i's probability of winning object m *strictly* increases as the m's component of the bid vector changes from b_{im} to $b_{im} + \epsilon$, then

$$v_{im} \leq -\frac{\left(P_{-i}(b_i|Z_{-i}) - P_{-i}(b_i + \epsilon e_m|Z_{-i})\right)^T \kappa(Z_i, W, X)}{\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im} + \epsilon|Z_{-i})} + \frac{b_{im}\Gamma_{-i}(b_{im}|Z_{-i}) - (b_{im} + \epsilon)\Gamma_{-i}(b_{im} + \epsilon|Z_{-i})}{\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im} + \epsilon|Z_{-i})}. \quad (15)$$

If $\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im} + \epsilon|Z_{-i}) = 0$, then $P_{-i}(b_i|Z_{-i}) - P_{-i}(b_i + \epsilon e_m|Z_{-i}) = 0$ and we obtain the following inequality that clearly holds:

$$0 \geq -\epsilon\Gamma_{-i}(b_{im}|Z_{-i}).$$

If there exists a known scalar $\bar{v} < \infty$ such that $V_{im} \leq \bar{v}$ with probability 1 for any m (note that $\bar{v}$ could be strictly outside the support of V_{im}), then in this situation we can just bound v_{im} from above by $\bar{v}$.

Analogously, taking

$$b = b_i - \epsilon e_m \in \mathcal{B}_i$$

for $\epsilon > 0$, obtain from (14) that

$$v_{im}\left(\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im}-\epsilon|Z_{-i})\right) + \left(P_{-i}(b_i|Z_{-i}) - P_{-i}(b_i+\epsilon e_m|Z_{-i})\right)^T \kappa(Z_i, W, X) \geq b_{im}\Gamma_{-i}(b_{im}|Z_{-i}) - (b_{im}-\epsilon)\Gamma_{-i}(b_{im}-\epsilon|Z_{-i}).$$

If bidder i's probability of winning object m *strictly* increases as the m's component of the bid vector changes from $b_{im}-\epsilon$ to b_{im}, then

$$v_{im} \geq -\frac{\left(P_{-i}(b_i|Z_{-i}) - P_{-i}(b_i+\epsilon e_m|Z_{-i})\right)^T \kappa(Z_i, W, X)}{\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im}-\epsilon|Z_{-i})} + \frac{b_{im}\Gamma_{-i}(b_{im}|Z_{-i}) - (b_{im}-\epsilon)\Gamma_{-i}(b_{im}-\epsilon|Z_{-i})}{\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im}-\epsilon|Z_{-i})}. \quad (16)$$

If $\Gamma_{-i,m}(b_{im}|Z_{-i}) - \Gamma_{-i,m}(b_{im}-\epsilon|Z_{-i}) = 0$, then $P_{-i}(b_i|Z_{-i}) - P_{-i}(b_i - \epsilon e_m|Z_{-i}) = 0$ and we obtain the inequality $0 \geq -\epsilon\Gamma_{-i}(b_{im}|Z_{-i})$, which implies that $\Gamma_{-i}(b_{im}|Z_{-i}) = 0$. If there exists a known scalar $\underline{v} \geq 0$ such that $V_{im} \geq \underline{v}$ with probability 1 for any m (note that $\underline{v}$ could be strictly outside the support of V_{im}), then in this situation we can just bound v_{im} from below by $\underline{v}$.

Inequalities (15) and (16) will be the basis for our analysis. But before we proceed let us introduce some notations. Let $\Delta^+_{\epsilon,m}[f(u)]$ and $\Delta^-_{\epsilon,m}[f(u)]$ denote differences in the values of $f(\cdot)$ associated with adding ϵ and $-\epsilon$ to the mth component of u respectively:

$$\Delta^+_{\epsilon,m}[f(u)] = f(u+\epsilon e_m) - f(u),$$
$$\Delta^-_{\epsilon,m}[f(u)] = f(u-\epsilon e_m) - f(u),$$

where e_m denotes the M_i-dimensional mth unit vector.

For each $b_i \in \mathcal{B}_i$, define $I^-_{\epsilon,m}(b_i|Z_{-i})$, $I^+_{\epsilon,m}(b_i|Z_{-i})$ as follows:

$$I^-_{\epsilon,m}(b_i|Z_{-i}) = \left\{\underline{v} \text{ if } \Delta^-_{\epsilon,m}[\Gamma_{-i}(b_i|Z_{-i})] = 0,\ \frac{\Delta^-_{\epsilon,m}[b_i^T\Gamma_{-i}(b_i|Z_{-i})]}{\Delta^-_{\epsilon,m}[\Gamma_{-i,m}(b_i|Z_{-i})]} \text{ else}\right\};$$
$$I^+_{\epsilon,m}(b_i|Z_{-i}) = \left\{\bar{v} \text{ if } \Delta^+_{\epsilon,m}[\Gamma_{-i}(b_i|Z_{-i})] = 0,\ \frac{\Delta^+_{\epsilon,m}[b_i^T\Gamma_{-i}(b_i|Z_{-i})]}{\Delta^+_{\epsilon,m}[\Gamma_{-i,m}(b_i|Z_{-i})]} \text{ else}\right\}.$$

Also, for each $b_i \in \mathcal{B}_i$, define the following $S^-_{\epsilon,m}(b_i|Z_{-i})$ and $S^+_{\epsilon,m}(b_i|Z_{-i})$:

$$S^-_{\epsilon,m}(b_i|Z_{-i}) = \left\{0 \text{ if } \Delta^-_{\epsilon,m}[\Gamma_{-i}(b_i|Z_{-i})] = 0,\ \frac{\Delta^-_{\epsilon,m}[P_{-i}(b_i|Z_{-i})]}{\Delta^-_{\epsilon,m}[\Gamma_{-i,m}(b_i|Z_{-i})]} \text{ else}\right\};$$

$$S^+_{\epsilon,m}(b_i|Z_{-i}) = \left\{0 \text{ if } \Delta^+_{\epsilon,m}[\Gamma_{-i}(b_i|Z_{-i})] = 0,\ \frac{\Delta^+_{\epsilon,m}[P_{-i}(b_i|Z_{-i})]}{\Delta^+_{\epsilon,m}[\Gamma_{-i,m}(b_i|Z_{-i})]} \text{ else}\right\}.$$

For $K \in \mathcal{K}_i$, let $\tilde{F}^-_{im}(\cdot|K; Z_{-i})$ denote the c.d.f. of

$$\sup_{\epsilon>0}\left(I^-_{\epsilon,m}(b_i|Z_{-i}) - S^-_{\epsilon,m}(b_i|Z_{-i})^T K\right),$$

and let $\tilde{F}^+_{im}(\cdot|K; Z_{-i})$ denote the c.d.f. of

$$\inf_{\epsilon>0}\left(I^+_{\epsilon,m}(b_i|Z_{-i}) - S^+_{\epsilon,m}(b_i|Z_{-i})^T K\right)$$

.

Then inequalities (15) and (16) imply that a superset of the identified set of $\kappa(Z_i, W, X)$ for bidder i can be found as

$$\bigcap_{m=1}^{M_i} \tilde{\mathcal{K}}_{i,m}(Z_i, W, X),$$

where $\tilde{\mathcal{K}}_{i,m}(Z_i, W, X)$ is defined as

$$\tilde{\mathcal{K}}_{i,m}(Z_i, W, X) = \{K \in \mathcal{K}_i \mid \tilde{F}^+_{im}(\cdot|K; Z_{-i}) \le \tilde{F}^-_{im}(\cdot|K; Z'_{-i}) \quad \forall\, Z_{-i}, Z'_{-i} \in \mathcal{Z}_{-i}|Z_i, W, X\}.$$

Let us denote this superset as $\mathcal{H}^{(1)}_{i,\kappa}(Z_i, W, X)$.

Now we can construct supersets of the identified sets for the distributions of standalone valuations. As $\mathcal{F}_c(\mathbb{R}^p)$ we denote the set of all continuous cumulative distribution functions on $\mathbb{R}^p$.

A superset of the identified set for the c.d.f. of the standalone valuation V_{im} conditional on Z_i, X can be found as the set of univariate functions $F_{im}(\cdot) \in \mathcal{F}_c(\mathbb{R})$ such that for any $\eta \in \mathbb{R}$,

$$F_{im}(\eta) \in \bigcap_{W \in \mathcal{W}|Z_i,X}\ \bigcap_{\kappa_0 \in \mathcal{H}^{(1)}_{i,\kappa}(Z_i,W,X)}\ \bigcap_{Z_{-i},Z'_{-i} \in \mathcal{Z}_{-i}|Z_i,W,X} [\tilde{F}^+_{im}(\eta|\kappa_0; Z_{-i}), \tilde{F}^-_{im}(\eta|\kappa_0; Z'_{-i})]\}. \quad (17)$$

Here we applied the exclusion restriction that the distribution of standalone valuations conditional on Z_i, W, X does not depend on W. Let us denote this superset as $\mathcal{H}^{(1)}_{i,F_m}(Z_i, X)$.

Our final step is to construct a superset $\mathcal{H}^{(1)}_{i,F}(Z_i, X)$ for the identified set for the joint distribution of the vector of standalone valuations. *$\mathcal{H}^{(1)}_{i,F}(Z_i, X)$ can be found as the set of M_i-variate functions $F_i(\cdot) \in \mathcal{F}_c(\mathbb{R}^{M_i})$ such that each mth marginal distribution function generated by $F_i(\cdot)$ belongs to $\mathcal{H}^{(1)}_{i,F_m}(Z_i, X)$, $m = 1, \ldots, M_i$. Moreover, for any $\eta = (\eta_1, \ldots, \eta_{M_i})$,*

$$F_i(\eta) \leq \min_{m=1,\ldots,M_i} \inf_{W\in\mathcal{W}|Z_i,X} \inf_{\kappa_0\in\mathcal{H}^{(1)}_{i,\kappa}(Z_i,W,X)} \inf_{Z_{-i}\in\mathcal{Z}_{-i}|Z_i,W,X} \tilde{F}^{+}_{im}(\eta_m|\kappa_0; Z_{-i}), \tag{18}$$

$$F_i(\eta) \geq \max\left\{\sum_{m=1}^{M_i} \sup_{W\in\mathcal{W}|Z_i,X} \sup_{\kappa_0\in\mathcal{H}^{(1)}_{i,\kappa}(Z_i,W,X)} \sup_{Z_{-i}\in\mathcal{Z}_{-i}|Z_i,W,X} \tilde{F}^{-}_{im}(\eta_m|\kappa_0; Z_{-i}) - M_i + 1, 0\right\}, \tag{19}$$

where we employed the well known result on sharp Frechet-Hoeffding bounds for joint distributions.

Below we provide an expectations version of the partial identification argument. Even though the supersets in the expectations approach will be larger than those discussed previously, computationally they are easier to obtain. Before describing these supersets, let us define $M_i \times 1$ vectors $\Psi^-_\epsilon(Z_{-i})$, $\Psi^+_\epsilon(Z_{-i})$ and $M_i \times 2^{M_i}$ matrices $\chi^-_\epsilon(Z_{-i})$, $\chi^+_\epsilon(Z_{-i})$ as follows:

$$\begin{aligned}
\Psi^-_\epsilon(Z_{-i}) &\equiv \left[E[I^-_{\epsilon,m}(B_i|Z_{-i})|Z_{-i}]\right]^{M_i}_{m=1}\\
\Psi^+_\epsilon(Z_{-i}) &\equiv \left[E[I^+_{\epsilon,m}(B_i|Z_{-i})|Z_{-i}]\right]^{M_i}_{m=1}\\
\chi^-_\epsilon(Z_{-i}) &\equiv \left[E[S^-_{\epsilon,m}(B_i|Z_{-i})|Z_{-i}]^T\right]^{M_i}_{m=1}\\
\chi^+_\epsilon(Z_{-i}) &\equiv \left[E[S^+_{\epsilon,m}(B_i|Z_{-i})|Z_{-i}]^T\right]^{M_i}_{m=1}.
\end{aligned}$$

Then, applying the expectation over the distribution of bids conditional on Z_i, W, X to inequalities (15) and (16) and pooling restrictions across Z_{-i}, Z'_{-i} and $m = 1, \ldots, M_i$, we establish that a superset of the identified set for $\kappa(Z_i, W, X)$ can be found in the following way:

$$\mathcal{H}^{(2)}_{i,\kappa}(Z_i, W, X) = \bigcap_{\epsilon>0} \hat{\mathcal{K}}^\epsilon_i(Z_i, W, X),$$

where $\hat{\mathcal{K}}^\epsilon_i(Z_i, W, X)$ is defined as

$$\begin{aligned}
\hat{\mathcal{K}}^\epsilon_i(Z_i, W, X) \equiv \Big\{K \in \mathcal{K}_i \;\Big|\; &\left(\Psi^-_\epsilon(Z_{-i}) - \Psi^+_\epsilon(Z'_{-i})\right)\\
&- \left(\chi^-_\epsilon(Z_{-i}) - \chi^+_\epsilon(Z'_{-i})\right) K \leq 0 \text{ for all } Z_{-i}, Z'_{-i} \in \mathcal{Z}_{-i}|Z_i, W, X\Big\}.
\end{aligned}$$

Notice two features of $\mathcal{H}^{(2)}_{i,\kappa}(Z_i, W, X)$. First, it can be represented as the intersection of a set of half-spaces in $\mathcal{K}_i$, where half-spaces are bounded by hyperplanes involving slope vectors $(\chi^{-}_{\epsilon,m}(Z_{-i}) - \chi^{+}_{\epsilon,m}(Z'_{-i}))$ and intercepts $(\Psi^{-}_{\epsilon,m}(Z_{-i}) - \Psi^{+}_{\epsilon,m}(Z'_{-i}))$, and the intersection is taken over collections of $(Z_{-i}, Z'_{-i}, \epsilon, m)$.

Second, if G_i is absolutely continuous, then $\mathcal{H}^{(2)}_{i,\kappa}(Z_i, W, X)$ is a singleton, and as we show below, the analysis of $\mathcal{H}^{(2)}_{i,\kappa}(Z_i, W, X)$ essentially becomes our identification strategy in the case of point identification. Indeed, bidder i's objective function is differentiable at almost every observed b_i. Hence as $\epsilon \to 0$ we will have for all m

$$\lim_{\epsilon\to 0} \frac{\Delta^{-}_{\epsilon,m} b_i \Gamma_{-i}(b_i|Z_{-i})}{\Delta^{-}_{\epsilon,m}\Gamma_{-i,m}(b_i|Z_{-i})} = \lim_{\epsilon\to 0} \frac{\Delta^{-}_{\epsilon,m} b_i \Gamma_{-i}(b_i|Z_{-i})/\epsilon}{\Delta^{-}_{\epsilon,m}\Gamma_{-i,m}(b_i|Z_{-i})/\epsilon} = \frac{\partial(b_i\Gamma_{-i}(b_i|Z_{-i}))/\partial b_{im}}{d\Gamma_{-i,m}(b_i|Z_{-i}))/db_{im}},$$

and therefore $\Psi^{-}_{\epsilon}(\cdot) \to \Psi(\cdot)$. Analogously, it is straightforward to show that $\Psi^{+}_{\epsilon}(\cdot) \to \Psi(\cdot)$, $\chi^{-}_{\epsilon} \to \chi(\cdot)$, and $\chi^{+}_{\epsilon} \to \chi(\cdot)$. Hence after applying the expectations operator, inequalities (15) and (16) imply that

$$\Psi(Z_{-i}) - \chi(Z_{-i})\kappa_0 \le \Psi(Z'_{-i}) - \chi(Z'_{-i})\kappa_0 \quad \forall\, Z_{-i}, Z'_{-i} \in \mathcal{Z}_{-i}|Z_i, W, X.$$

Noting that Z_{-i}, Z'_{-i} are interchangeable, we thus have for any $Z_{-i}, Z'_{-i} \in \mathcal{Z}_{-i}|Z_i, W, X$:

$$\begin{aligned} \Psi(Z_{-i}) - \chi(Z_{-i})\kappa_0 &\le \Psi(Z'_{-i}) - \chi(Z'_{-i})\kappa_0 \\ \Psi(Z'_{-i}) - \chi(Z'_{-i})\kappa_0 &\le \Psi(Z_{-i}) - \chi(Z_{-i})\kappa_0, \end{aligned}$$

or equivalently

$$\Psi(Z_{-i}) - \chi(Z_{-i})\kappa_0 = \Psi(Z'_{-i}) - \chi(Z'_{-i})\kappa_0 \quad \forall\, Z_{-i}, Z'_{-i} \in \mathcal{Z}_{-i}|Z_i, W, X.$$

But this is exactly the identification restriction invoked in Proposition 3 in the main text. Thus we can strictly generalize our existing identification results (which depend on differentiability a.e.) to partial identification for arbitrary G_i.

A superset for the identification set of the c.d.f. of V_{im} can be found as in (17) by replacing $\mathcal{H}^{(1)}_{i,\kappa}(Z_i, W, X)$ with $\mathcal{H}^{(2)}_{i,\kappa}(Z_i, W, X)$. Similarly, a superset for the identification set of the c.d.f. of vector V_i can be found as in (18) and (19) by replacing $\mathcal{H}^{(1)}_{i,\kappa}(Z_i, W, X)$ with $\mathcal{H}^{(2)}_{i,\kappa}(Z_i, W, X)$.

Appendix D: Complementarities depending on V

In this appendix, we explore prospects for generalizing our non-parametric identification results to the case where complementarities are additively separable functions of standalone valuations. In other words, conditional on Z, W, X the compementarities are stochastic but their randomness can be fully explained by the standalone valuations. As a special case, we consider a scenario when these functions are affine in standalone valuations. Such a case could arise if, for instance, winning two auctions together increases i's valuation for one or both objects by a fixed percentage.

Notation and definitions We say complementarities are *additively separable in standalone valuations* if for each ω that contains at least two non-zero components (that is, $\|\omega\|^2 \geq 2$), the complementarity for outcome ω is a function of the vector of standalone valuations $v_i = (v_{i1}, v_{i2}, \ldots, v_{iM_i})^T$ such that

$$K^\omega(v_i, Z_i, W, X) = \sum_{l:\,\omega_l = 1} \phi_l(v_{il}, Z_i, X, W) + \bar{K}^\omega(Z_i, X, W) \tag{20}$$

for some functions ϕ_l, $l = 1, \ldots, L$. If each function ϕ_l is linear in v_{il}, then we obtain the special case of *complementarities affine in v_i*:

$$K^\omega(v_i, Z_i, W, X) = \sum_{l:\,\omega_l = 1} \delta^l(Z_i, W, X) v_{il} + \bar{K}^\omega(Z_i, W, X), \quad \text{ if } \|\omega\|^2 \geq 2. \tag{21}$$

As usual, if ω contains at most one component equal to one (that is, $\|\omega\|^2 \leq 1$), then we set $K^\omega(v_i, Z_i, W, X) \equiv 0$. An interesting special case of (21) is when all δ^l are identical and $\bar{K}^\omega = 0$ for any ω. This case describes the situation of a constant relative complementarity – that is, when $K^\omega(v_i, Z_i, W, X)$ is a constant ratio of the additive valuation.

Now assume that complementarities are affine in v_i, and define an $M_i \times 1$ vector $\delta(Z_i, W, X)$ and an $M_i \times M_i$ matrix $D(\delta(Z_i, W, X))$ as follows:

$$\begin{aligned}
\delta(Z_i, W, X) &\equiv (\delta^1(Z_i, W, X), \delta^2(Z_i, W, X), \ldots, \delta^{M_i}(Z_i, W, X))^T \\
D(\delta(Z_i, W, X)) &\equiv \operatorname{diag}(\delta^1(Z_i, W, X), \delta^2(Z_i, W, X), \ldots, \delta^{M_i}(Z_i, W, X)).
\end{aligned}$$

To write this in a convenient vector-matrix notation, let A_i denote the $2^{M_i} \times 2^{M_i}$ matrix such that its submatrix $(a_{lj})_{l,j = M_i+2, \ldots, 2^M}$ coincides with the identity matrix of size $2^{M_i} - M_i - 1$, with all the other elements of A_i being 0. We then have

$$K(v_i, Z_i, W, X) = A_i \Omega_i D(\delta(Z_i, W, X)) v_i + \bar{K}(Z_i, W, X),$$

where $\bar{K}(Z_i, W, X)$ denotes the $2^{M_i} \times 1$ vector of constant components in the complementarities (obviously, $\bar{K}(Z_i, W, X) \in \mathcal{K}_i$). Clearly, the rank of matrix $A_i \Omega_i$ is equal to M_i.

As can be seen, the functional form of complementarities does not depend on Z_{-i}. As we show below, under weak conditions there is enough variation in $Z_{-i} \,|\, Z_i, W, X$ to determine the linear (in v_{il}) part of complementarities as well as the constant part.

Non-parametric identification Using the first-order conditions and taking into account the form of $K(v_i, Z_i, W, X)$, obtain

$$v_i = b_i + [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \Gamma_{-i}(b_i|Z_{-i}) - [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \nabla_b P_{-i}(b_i|Z_{-i})^T \left[A_i \Omega_i D(\delta) v_i + \bar{K}\right],$$

where for notational simplicity conditioning on Z_i, W, X is omitted from the notation in the rest of this Appendix. Rewrite that system of equations by collecting all terms with v_i on the left-hand side:

$$\begin{aligned}\left(I_{M_i} + [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \nabla_b P_{-i}(b_i|Z_{-i})^T A_i \Omega_i D(\delta)\right) v_i = b_i + [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \Gamma_{-i}(b_i|Z_{-i}) \\ - [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \nabla_b P_{-i}(b_i|Z_{-i})^T \bar{K},\end{aligned}$$

and introduce a notation for the matrix in front of v_i on the left-hand side:

$$\Pi(b_i, \delta, Z_{-i}) \equiv I_{M_i} + [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \nabla_b P_{-i}(b_i|Z_{-i})^T A_i \Omega_i D(\delta).$$

Define $\Delta(Z_{-i})$ as the set of $\delta \in \Re^{M_i}$ such that

$$\Pi(b_i, \delta, Z_{-i}) \text{ is non-singular for almost all } b_i.$$

This set is non-empty as e.g. $0 \in \Delta(Z_{-i})$. If $\delta \in \Delta(Z_{-i})$, then we can multiply the system from the left by $\Pi(b_i, \delta, Z_{-i})^{-1}$ resulting in

$$\begin{aligned} v_i = \Pi(b_i, \delta, Z_{-i})^{-1} b_i + \Pi(b_i, \delta, Z_{-i})^{-1} [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \Gamma_{-i}(b_i|Z_{-i}) \\ - \Pi(b_i, \delta, Z_{-i})^{-1} [\nabla_b \Gamma_{-i}(b_i|Z_{-i})]^{-1} \nabla_b P_{-i}(b_i|Z_{-i})^T \bar{K}.\end{aligned}$$

Assuming that $\delta \in \Delta(Z_{-i})$ and carrying on with fixed Z_i, W, X, let us denote

$$\begin{aligned}\mathcal{D}_1(\delta, Z_{-i}) &\equiv E_{B_i}\left[\Pi(B_i, \delta, Z_{-i})^{-1} B_i \big| Z_{-i}\right] + E_{B_i}\left[\Pi(B_i, \delta, Z_{-i})^{-1} [\nabla_b \Gamma_{-i}(B_i|Z_{-i})]^{-1} \Gamma_{-i}(B_i|Z_{-i}) \big| Z_{-i}\right], \\ \mathcal{D}_2(\delta, Z_{-i}) &\equiv E_{B_i}\left[\Pi(B_i, \delta, Z_{-i})^{-1} [\nabla_b \Gamma_{-i}(B_i|Z_{-i})]^{-1} \nabla_b P_{-i}(B_i|Z_{-i})^T \big| Z_{-i}\right].\end{aligned}$$

Keeping Z_i, W, X fixed, let us draw another value Z'_{-i} from $\mathcal{Z}_{-i}|Z_i, W, X$. Due to the assumptions made on the distribution of the standalone valuations, $E[V_i|Z_i, Z_{-i}, W, X] = E[V_i|Z_i, Z'_{-i}, W, X]$. Therefore, for $\delta \in \Delta(Z_{-i}) \cap \Delta(Z'_{-i})$,

$$\mathcal{D}_1(\delta, Z'_{-i}) - \mathcal{D}_1(\delta, Z_{-i}) = \left(\mathcal{D}_2(\delta, Z'_{-i}) - \mathcal{D}_2(\delta, Z_{-i})\right) \bar{K}.$$

For fixed Z_i, W, X, this system has $2^{M_i} - 1$ unknowns (M_i in δ and $2^{M_i} - M_i - 1$ in $\bar{K}$) and M_i equations. This gives us the following result.

Proposition 4. *Suppose that for $(Z_i, W, X) \in \mathcal{Z}_i \times \mathcal{W} \times \mathcal{X}$, there exist $J + 1 \geq (2^{M_i} - 1)/M_i + 1$ vectors $Z_{-i,0}, Z_{-i,1}, \ldots, Z_{-i,J}$ in the support $\mathcal{Z}_{-i}|Z_i, W, X$ such that there is a* *unique* $\delta \in \bigcap_{j=0}^{J} \Delta(Z_{-i,j})$ *and a unique* $\kappa \in \mathcal{K}_i$ *that solve the system of* $J \cdot M_i$ *equations*

$$\mathcal{D}_1(\delta, Z_{-i,j}) - \mathcal{D}_1(\delta, Z_{-i,0}) = (\mathcal{D}_2(\delta, Z_{-i,j}) - \mathcal{D}_2(\delta, Z_{-i,0}))\,\kappa, \quad j = 1, \ldots, J. \tag{22}$$

Then the values of $\delta(Z_i, W, X)$ *and* $\bar{K}(Z_i, W, X)$ *are identified, and thus, the complementarity function is identified for these values of* Z_i, W, X.

System (22) is non-linear in δ. However, for each fixed $\delta \in \bigcap_{j=0}^{J} \Delta(Z_{-i,j})$, this system is linear in κ. Proposition 4 implies that in the case of identification it is not possible to have a situation when for different δ_1 and δ_2, where $\delta_1, \delta_2 \in \bigcap_{j=0}^{J} \Delta(Z_{-i,j})$, system (22) has solutions $\kappa_1 \in \mathcal{K}_i$ and $\kappa_2 \in \mathcal{K}_i$, respectively. Thus, in this sense the question of identification of $\delta(Z_i, W, X)$ and $\bar{K}(Z_i, W, X)$ comes down to the question of the *existence* of a solution to a system of linear equations: (22) can have a solution κ for one δ only, and for that δ it has to be unique. Using the Kronecker-Capelli theorem, which gives the necessary and sufficient conditions for the existence of a solution to a system of linear equations, and also the necessary and sufficient conditions for the uniqueness of such a solution, we formulate the identification result in the Proposition 5 below.

Before we proceed to Proposition 5, let is rewrite (22) in a more convenient way. At the moment k has to satisfy certain restrictions (namely, the first $M_i + 1$ components of this vector are 0) and we first want to rewrite it through an unrestricted parameter to apply certain tools from algebra. Let E_i denote the $2^{M_i} \times (2^{M_i} - M_i - 1)$ matrix such that its submatrix $(\tilde{e}_{ij})_{i=M_i+2,\ldots,2^{M_i},\, j=1,\ldots,2^{M_i}-l_i-1}$ coincides with the identity matrix of size $2^{M_i} - M_i - 1$, and all its other elements (that is, all the elements in the first $M_i + 1$ rows) are equal to zero. For every $\kappa \in \mathcal{K}_i$ there is a unique $\breve{\kappa} \in \mathbb{R}^{2^{M_i} - M_i - 1}$ such that

$$\kappa = E_i \breve{\kappa}.$$

Obviously, this $\breve{\kappa}$ is a parameter that does not have to satisfy any prior restrictions. It is formed by the last $2^{M_i} - M_i - 1$ values in κ. System (22) can equivalently be written as

$$\mathcal{D}_1(\delta, Z_{-i,j}) - \mathcal{D}_1(\delta, Z_{-i,0}) = (\mathcal{D}_2(\delta, Z_{-i,j})E_i - \mathcal{D}_2(\delta, Z_{-i,0})E_i)\ \breve{\kappa}, \qquad j = 1, \ldots, J, \tag{23}$$

with $\breve{\kappa} \in \mathbb{R}^{2^{M_i} - M_i - 1}$. For a fixed δ, system (22) is linear in κ, has the $J \cdot M_i \times 2^{M_i}$ matrix of coefficients, and imposes restrictions on the solution κ by requiring that $\kappa \in \mathcal{K}_i$. Its equivalent representation (23) is linear in $\breve{\kappa}$ for a fixed δ, has the $J \cdot M_i \times (2^{M_i} - M_i - 1)$ matrix of coefficients, and does not impose any restrictions on the solution $\breve{\kappa} \in \mathbb{R}^{2^{M_i} - M_i - 1}$. This allows us to apply the Kronecker-Capelli theorem to system (23) in a straightforward way.

Proposition 5. *Suppose that for* $(Z_i, W, X) \in \mathcal{Z}_i \times \mathcal{W} \times \mathcal{X}$, *there exist* $J + 1 \geq (2^{M_i} - 1)/M_i + 1$ *vectors* $Z_{-i,0}, Z_{-i,1}, \ldots, Z_{-i,J}$ *in the support* $\mathcal{Z}_{-i}|Z_i, W, X$ *such that there is a unique* $\delta \in \bigcap_{j=0}^{J} \Delta(Z_{-i,j})$ *that satisfies the following two conditions:*

1. First,

$$rank\left([\mathbf{M}_1(\delta)\,|\,\mathbf{M}_2(\delta)]\right) = rank\left(\mathbf{M}_2(\delta)\right), \tag{24}$$

where $\mathbf{M}_2(\delta)$ denotes the $J \cdot M_i \times (2^{M_i} - M_i - 1)$ matrix

$$\mathbf{M}_2(\delta) \equiv \begin{bmatrix} \mathcal{D}_2(\delta, Z_{-i,1})E_i - \mathcal{D}_2(\delta, Z_{-i,0})E_i \\ \vdots \\ \mathcal{D}_2(\delta, Z_{-i,J})E_i - \mathcal{D}_2(\delta, Z_{-i,0})E_i \end{bmatrix},$$

and $\mathbf{M}_1(\delta)$ denotes the $J \cdot M_i \times 1$ vector

$$\mathbf{M}_1(\delta) \equiv \begin{bmatrix} \mathcal{D}_1(\delta, Z_{-i,1}) - \mathcal{D}_1(\delta, Z_{-i,0}) \\ \vdots \\ \mathcal{D}_1(\delta, Z_{-i,J}) - \mathcal{D}_1(\delta, Z_{-i,0}) \end{bmatrix}.$$

2. Moreover, this δ is such that $\mathbf{M}_2(\delta)$ has full column rank:

$$rank\left(\mathbf{M}_2(\delta)\right) = 2^{M_i} - M_i - 1. \tag{25}$$

Then the values of $\delta(Z_i, W, X)$ and $\bar{K}(Z_i, W, X)$ are identified, and thus, the complementarity function is identified for these values of Z_i, W, X.

Condition (24) requires that in system (23), the rank of the matrix of coefficients $\mathbf{M}_2(\delta)$ is equal to the rank of the augmented matrix $[\mathbf{M}_1(\delta)\,|\,\mathbf{M}_2(\delta)]$ for one δ only. The Kronecker-Capelli theorem guarantees then that (23) has a solution $\breve{\kappa}$ for that δ only. Condition (25) then guarantees this $\breve{\kappa}$ is determined uniquely, and, thus, $\kappa = E_i\breve{\kappa}$ is determined uniquely.

Note that all the identification conditions in Proposition 5 are formulated in terms of δ. The closed form for $\delta(Z_i, W, X)$ cannot be found but in practice one can find $\delta(Z_i, W, X)$ and $\bar{K}(Z_i, W, X)$ by solving, e.g., the following optimization problem:

$$\min_{\delta \in \bigcap_{j=0}^{J} \Delta(Z_{-i,j}),\, \breve{\kappa} \in \mathbb{R}^{2^{M_i} - M_i - 1}} Q(\delta, \breve{\kappa}, Z_i, W, X),$$

where

$$Q(\delta, \breve{\kappa}, Z_i, W, X) \equiv \left(\mathbf{M}_1(\delta) - \mathbf{M}_2(\delta)\breve{\kappa}\right)^T \left(\mathbf{M}_1(\delta) - \mathbf{M}_2(\delta)\breve{\kappa}\right).$$