Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Measuring the Performance of Large-Scale Combinatorial Auctions: A Structural Estimation Approach

Sang Won Kim, Marcelo Olivares, Gabriel Y. Weintraub

MANAGEMENT SCIENCE
Vol. 60, No. 5, May 2014, pp. 1180–1201
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)
informs
http://dx.doi.org/10.1287/mnsc.2013.1814

Measuring the Performance of Large-Scale Combinatorial Auctions: A Structural Estimation Approach

Sang Won Kim
CUHK Business School, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, skim@cuhk.edu.hk

Marcelo Olivares
Columbia Business School, Columbia University, New York, New York 10027; and University of Chile, Santiago, Chile, molivares@columbia.edu

Gabriel Y. Weintraub
Columbia Business School, Columbia University, New York, New York 10027, gweintraub@columbia.edu

The main advantage of a procurement combinatorial auction (CA) is that it allows suppliers to express cost synergies through package bids. However, bidders can also strategically take advantage of this flexibility, by discounting package bids and "inflating" bid prices for single items, even in the absence of cost synergies; the latter behavior can hurt the performance of the auction. It is an empirical question whether allowing package bids and running a CA improves performance in a given setting. In this paper, we develop a structural estimation approach that estimates the firms' cost structure using bidding data and use these estimates to evaluate the performance of the auction. To overcome the computational difficulties arising from the large number of bids observed in large-scale CAs, we propose a novel simplified model of bidders' behavior based on pricing package characteristics. We apply our method to the Chilean school meals auction, in which the government procures half a billion dollars' worth of meal services every year and bidders submit thousands of package bids. Our estimates suggest that bidders' cost synergies are economically significant in this application (∼5%), and the current CA mechanism achieves high allocative efficiency (∼98%) and reasonable margins for the bidders (∼5%). Overall, this work develops the first practical tool to evaluate the performance of large-scale first-price CAs commonly used in procurement settings.

Keywords: combinatorial auctions; procurement; empirical; structural estimation; auction design; public sector applications
History: Received May 2, 2012; accepted July 3, 2013, by Serguei Netessine, operations management. Published online in *Articles in Advance* January 6, 2014.

1. Introduction

In many important procurement settings, suppliers face cost synergies; for example, transportation service providers can lower costs by coordinating multiple deliveries in the same route, and producers can lower average costs by spreading a fixed cost across several units. Motivated by these types of settings, auction mechanisms that allow bidders to submit package bids for multiple units so that they can express their synergies have received much recent attention in practice and the academic literature. Indeed, these multiunit auctions, typically referred to as *combinatorial auctions* (CAs), have been implemented in many procurement applications. For example, Elmaghraby and Keskinocak (2004), Sandholm et al. (2006), and Hohner et al. (2003) describe applications at the Home Depot, Procter & Gamble, and Mars Inc., respectively. These types of auctions have also been implemented in nonprocurement settings, most notably in the auctions for wireless spectrum run by the Federal Communications Commission (FCC) (McDuff 2003).[1]

A central auction design question in multiunit settings is how allowing bidders to submit bids for packages of units impacts the performance of the mechanism. From the perspective of an auction designer, there are typically two measures that are relevant when evaluating performance: (1) efficiency, which compares the actual bidders' costs realized in the auction allocation relative to the minimum possible cost allocation that can be achieved; and (2) optimality, which relates to the total payment to the bidders by the auctioneer. The above design question is crucial because allowing for package bidding via a

[1] Cramton et al. (2006) provide an overview on CAs.

CA can have countering effects on the performance under these two measures, as we describe next.

On one hand, allowing package bids can enhance the performance especially in the presence of cost synergies. In many procurement applications, such as the examples mentioned above, bidders may have cost synergies due to *economies of scale*, which depend on the volume allocated to a given supplier, and *economies of density*, which depend on the proximity of the units in an allocation. If bidders were allowed only to submit bids for each unit separately, they would face the risk of winning some units but not others. This phenomenon, known as the *exposure problem*, may induce bidders to be less aggressive in expressing the economies of scale and density that arise from supplying multiple units. Enabling package bidding through a CA eliminates this risk, potentially leading to more efficient outcomes and lower procurement costs.

However, allowing package bids could also hurt the performance. As pointed out by Cantillon and Pesendorfer (2006) and Olivares et al. (2012), with a first-price rule, bidders can engage in strategic bundling in which they submit package discounts even in the absence of cost synergies. One motivation to do so may be to leverage a relative cost advantage in a unit (for which the bidder is the cost-efficient provider) into another unit (for which the bidder is not the efficient provider). The firm may attempt to win both units by submitting a "discounted" package bid for the bundle and "inflating" both single-unit bids. If the bidder wins the package, it will lead to an inefficient allocation in which a unit is not served by the lowest-cost supplier. In addition, package bidding can also lead to the *threshold problem*, in which "local" suppliers bidding for small packages free ride on each other to outbid "global" suppliers submitting bids on larger packages; this free riding can lead to less competitive bidding, higher margins, and thereby higher payments for the auctioneer. Milgrom (2000) and Baranov (2010) provide examples of the threshold problem.

Given the aforementioned trade-off, we expect that a CA should enhance the performance, relative to auction mechanisms that preclude package bidding, if cost synergies are strong and the incentives for the types of strategic behavior mentioned above are weak. However, analyzing the actual performance of a CA requires evaluating cost efficiency and the margins of the winning bidders, which is typically private and sensitive information of the bidders. Moreover, existing theory is not conclusive on how large the incentives for strategizing are in a specific application. Thus motivated, in this paper we develop and apply an empirical approach to evaluate the performance of first-price CAs based on observed bid data, and use it to inform the auction design.

A reduced-form analysis of the bid data like the one in Olivares et al. (2012) can be used to provide a direct measurement of the package discounts relative to single-unit bids observed in a CA. However, the presence of package discounts is not conclusive about the performance of the auction. Bid discounts may reflect cost synergies, but they could also reflect the types of strategic behavior alluded to above; bidders could inflate their single-unit bids relative to package bids to increase the probability of winning larger packages with relatively high margins, even in the absence of cost synergies. These *strategic markup reductions* also result in package discounts and a reduced-form analysis of the bid data cannot directly distinguish between this and a cost synergy-based explanation. This is limiting when evaluating the efficiency of the auction, since we expect a CA to perform well only if package bid discounts are mostly explained by cost synergies. Moreover, since a reduced-form analysis does not identify bidders' cost information from the observed bids, it cannot be used to evaluate alternative mechanism designs.

To overcome these limitations, we introduce a structural estimation approach that imposes a model of bidder's behavior to estimate bidders' supplying costs, and therefore disentangle cost synergies and strategic markup reductions from the observed bid discounts. Our method is based on the influential work of Guerre et al. (2000) for single-unit auctions that was later extended by Cantillon and Pesendorfer (2006) and was applied to the London bus routes CAs with two or three units. More specifically, Cantillon and Pesendorfer (2006) conduct the structural estimation of first-price CAs in two steps. In the first step, a statistical distribution of the competitors' bids is estimated from bidding data. In the second step, the first-order conditions from the bidder's profit maximization problem are used to find the imputed costs that would rationalize the bids observed in the data. These first-order conditions involve beliefs about the competitors' bidding behavior, and the distribution estimated in the first step is used to sample competitors' bids and form these beliefs. The estimated costs enable the calculation of the cost efficiency and bidders' margins in the CA to evaluate its performance. They also allow to evaluate alternative auction designs. We show, however, that this approach cannot be directly applied to large-scale CAs with many units, because of the high dimensionality of the bid vectors. This is an important limitation for many real-world procurement applications of CAs; for example, Caplice and Sheffi (2006) and Bichler et al. (2006) report CAs for transportation and procurement of inputs that typically involve hundreds of units.

Consequently, an important methodological contribution of our work is the development of a novel approach to apply structural estimation to large-scale first-price CAs. We introduce a "simplified" version of the bidder's problem where the markups charged on package bids are chosen based on a reduced set of package characteristics. With this simplification, the first-order conditions of the bidder's problem become computationally and econometrically tractable. We impose reasonable restrictions to the structure of the markups that reduce the complexity of the bidders problem but still provide sufficient flexibility to capture strategic behavior that can hurt the performance of a CA. In addition, we introduce a parsimonious, yet flexible parametric description of the distribution of competitors' bids for CAs that involve heterogeneous units, and scale and density discounts. This specification makes the estimation of the distribution of competitors' bids tractable. Overall, these two simplifications make the structural estimation feasible in large-scale CAs.

We apply our method to the Chilean school meals CA in which the government procures half a billion dollars' worth of meal services every year to feed 2.5 million children daily. This is one of the largest and most important social programs run by the Chilean government. The application fits well within the class of large-scale CAs: each auction has about 30 units, and firms submit hundreds of bids (for more details on the auction, see Epstein et al. 2002). This application serves as a template to illustrate how to apply our method and shows how its results can provide managerial insights into the auction design. In particular, the government officials running this auction have considered revising its format and we use the structural approach developed here to inform this question.

Our results show that cost synergies are significant. Package discounts can be as large as 6% of the average bid price, and our estimates show that most of the discounts (79%–86%) represent cost synergies, which amount to 5% of the average cost. The rest of the bid discounts are explained by strategic markup reductions, and we also use the estimation results to pinpoint the package bids for which bidders engage in strategic bundling. Overall, the estimated costs reveal that the CA achieves a high efficiency (in the order of 98%–99%) and reasonable margins for the bidders (in the order of 4%–5%). Drivers of this result are the relatively large cost synergies and the high level of competition in the auction; there is a reasonable number of firms and most firms compete in all units and submit many package bids. For this reason, firms do not seem to have enough market power to significantly harm efficiency by using the flexibility that package bidding provides with strategic motivations. In summary, our results suggest that allowing package bidding and running a CA seems appropriate in this setting. Further, to the best of our knowledge, this is the first paper to show that a CA performs well in a real-world application. Finally, we also use the cost estimates to evaluate the performance of a Vickrey-Clarke-Groves mechanism, showing that it achieves a reasonable total procurement cost, which counters recent theoretical results.

The structural estimation approach is prevalent in the economics literature and has been used to study multiunit auctions, such as wholesale electricity auctions (Reguant 2011), FCC spectrum auctions (Fox and Bajari 2013), and treasury auctions (Hortaçsu and McAdams 2010, Kastl 2011). Athey and Haile (2006), Hendricks and Porter (2007), and Paarsch and Hong (2006) provide comprehensive surveys about structural estimation of auction models. To the best of our knowledge, we are the first to develop a structural approach to tackle large-scale CAs and test it with a real application. Our work is also related to the growing literature in operations management that uses structural estimation; for example, it has been used to estimate customer waiting costs (Allon et al. 2011, Akşin et al. 2013), overage/underage costs of a newsvendor (Olivares et al. 2008), and consumer forward-looking behavior in the airline industry (Li et al. 2014). We add to this stream of research by applying structural estimation in a service procurement setting, an important area in operations and supply chain management where this approach has not been used.

2. Structural Estimation Approach for Combinatorial Auctions

This section develops a structural estimation framework to estimate the primitives of first-price single-round sealed-bid CAs. The standard structural approach to estimate auctions was pioneered by Guerre et al. (2000) (hereafter, GPV) for single-unit auctions. Cantillon and Pesendorfer (2006) (hereafter, CP) extended this approach to CAs and applied it to the London bus route auctions with three or fewer units. The structural approach introduced below closely follows the approach in CP, with some differences that we specify. In addition, at the end of this section, we discuss the limitations of applying this approach to large-scale CAs, because of the high dimensionality of the bid vectors.

First, we describe the basic setting of a CA. Let U denote the set of N units to be procured by an auctioneer. There is a set F of supplier firms, referred to as bidders and indexed by f. A package or combination, indexed by a, is a nonempty subset of units in U. We let $\mathcal{A}$ denote the set of all possible packages and

$A = |\mathcal{A}| (= 2^N - 1)$ be the total number of them. Let b_{af} denote the bid price asked by bidder f to supply package a, and $b_f = \{b_{af}\}_{a\in\mathcal{A}}$ the *bid vector* containing all bids from that bidder.

The following assumption describes the auction format.

Assumption 1 (Auction Format). *The auction has a first-price single-round sealed-bid format, so that bidders submit their bids simultaneously and winning bidders are paid their submitted bid prices for the packages awarded to them. The auction mechanism determines the winning bids by solving the following mathematical integer program, referred to as the winner determination problem*:

$$\begin{aligned} &\text{minimize} \quad \sum_{a\in\mathcal{A},\, f\in F} b_{af} x_{af} \\ &\text{subject to} \quad x \in X, \quad x_{af} = \{0,1\}, \quad \forall\, a\in\mathcal{A},\ f\in F, \end{aligned} \tag{1}$$

where x_{af} is a binary decision variable that is equal to one if and only if package a is assigned to bidder f, and $x = \{x_{af}\}_{a\in\mathcal{A},\, f\in F}$. We denote by X the set of feasible allocations; the set imposes that each unit is allocated to exactly one bidder, that each bidder can win at most one package, and potentially some additional allocative constraints.

The winner determination problem (1) minimizes the total procurement cost of the auctioneer, given the submitted bids. We note that the additional constraints in the set of feasible allocations could impose, for example, market share constraints that limit the maximum package size that a single bidder can be awarded, which may be used to keep a diversified supplier base. In §5 and the online appendix (available at http://www.columbia.edu/~gyw2105), we provide more details on the winner determination problem and its integer program formulation in the context of our specific empirical application.

The structural estimation approach is based on an auction model with private information and requires assumptions on the bidders' information structure and bidding behavior in order to identify their costs.

Assumption 2 (Bidders' Costs). *Bidders have independent private costs. In particular, given an auction, each bidder $f \in F$ gets an independent random draw of a cost vector $c_f = \{c_{af}\}_{a\in\mathcal{A}}$, in which c_{af} is the cost of supplying package a for bidder f.*

Before submitting its bid, each bidder observes its own vector of costs, but does not observe the costs' realizations of its competitors. Moreover, because costs are private, a bidder's costs only depend on its own private signal and it is not a function of the costs' realizations of other bidders. Based on this information structure, we make the following assumption on the bidders' bidding strategies.

Assumption 3 (Strategies). *Bidders are risk neutral and play pure bidding strategies. In particular, for a given auction, a bidder's strategy is a function $b_f\colon \Re_+^A \mapsto \Re_+^A$ that depends on its own costs c_f. Bidders place bids on all possible combinations of units.*

In our sealed-bid format, bidders submit their bids in a game of incomplete information without directly observing the bids nor the cost realizations of their competitors. Therefore, bidders face uncertainty on whether they will win any given package. For each bidder, we capture this uncertainty with the vector $G_f(b_f) = \{G_{af}(b_f)\}_{a\in\mathcal{A}}$, where $G_{af}(b_f)$ is the probability that bidder f wins package a with bid vector b_f. Using vector notation, we can then write a bidder's expected profit maximization problem as

$$\underset{b\in\Re_+^A}{\text{maximize}}\, (b-c)^T G(b), \tag{2}$$

where v^T denotes the transpose of a vector v. Note that each bidder has its own optimization problem with its own cost and winning probability vectors. To simplify the notation, we omit the subscript f whenever clear from the context.

To formulate the optimization problem above, a bidder needs to form expectations about the bidding behaviors of its competitors, so that it can evaluate the vector of winning probabilities $G(b)$, for a given value of b. Note that if bidder f anticipates that bidder f' uses a bidding strategy $b_{f'}(\cdot)$, bidder f''s bids are random from bidder f's perspective; they correspond to the composition $b_{f'}(c_{f'})$, where $c_{f'}$ is the cost vector for bidder f'. Note that $c_{f'}$ is random from bidder f's perspective, because it is private information. Assumption 4, described next, formalizes this. Assumptions 1–4 are kept throughout the paper, and we discuss their validity in the context of our application in §5.1.

Assumption 4 (Bid Distributions). (a) *Consider a given auction and any bidder $f \in F$. From the perspective of other bidders, the bid vector of firm f, $b_f = b_f(c_f)$, is random and is given by the composition of the strategy used by firm f in the auction and its random cost vector c_f (see Assumptions* 2 *and* 3*). Accordingly, denote by $H_f(\cdot \mid Z)$ the distribution of b_f, where Z is a vector of observable bidders and auction characteristics. This distribution is common knowledge among bidders.*

(b) *For all bidders $f \in F$, the competitors' random bid vectors $\{b_{f'}\}_{f'\neq f}$ are mutually independent conditional on Z.*

(c) *For all bidders $f \in F$, the distributions of competitors' bids $\{H_{f'}(\cdot \mid Z)\}_{f'\neq f}$ and the winner determination problem* (1) *induce the beliefs on the winning probabilities $G_f(b_f)$, for any given b_f.*

(d) *For all bidders $f \in F$, $H_f(\cdot \mid Z)$ has a continuous density everywhere.*

We note that whereas Assumptions 1–3 (or similar variations of them) are commonly made in the literature, Assumption 4 departs from the standard structural approach followed by CP and GPV in the following sense. The standard approach assumes that the primitives of the model such as the number of bidders, the probability distribution of costs, and the utility functions are common knowledge and that bidders play a Bayes Nash equilibrium (BNE) of the game induced by the auction. In many settings, such as the first-price single-unit auction studied in GPV, this is well justified because under mild conditions a unique symmetric pure strategy BNE always exists. However, there is no theoretical result available that guarantees existence of a *pure strategy* BNE in a CA. As we describe next, Assumption 4 is weaker than assuming BNE play, but still lends itself to using the two-step estimation approach in CP.[2]

More specifically, note that assuming pure strategy BNE play imposes two conditions for each bidder: (i) the bidder correctly anticipates the strategies of its competitors, and therefore correctly estimates the vector of winning probabilities given its own bids; and (ii) the bidder selects a bid vector that maximizes its expected profit, given its costs and the winning probabilities function. Although conditions (a)–(c) in Assumption 4 are weaker than condition (i) in a private cost setting, they impose the same restriction over bidders' beliefs that we use in our structural estimation approach: bidders in the auction can correctly anticipate their winning probabilities. This follows because Assumption 4 imposes that bidders' beliefs on winning probabilities are induced by the distributions $\{H_f(\cdot \mid Z)\}_{f\in F}$, which are constructed with the correct strategies used by the competitors in the auction together with their actual costs' distributions. We also make a weaker assumption relative to the aforementioned condition (ii) imposed by BNE: we will only assume that each bidder selects a bid vector that satisfies the necessary first-order conditions of the expected profit maximization problem (2). Despite these differences in the formulation of the structural model, the first-order conditions introduced below in §2.1 are the same as the ones used by CP to identify bidders' costs.

Condition (d) in Assumption 4 guarantees the differentiability of the winning probability vector $G(\cdot)$ that is needed to use the first-order conditions for estimation. Note that this assumption is over the bids' distributions, which are endogenously determined in the auction game. Although we would prefer to make assumptions over model primitives that imply the assumptions on behavior, the lack of theoretical results regarding the existence and characterization of pure strategy equilibria in CAs does not allow us to follow this approach. We formalize the differentiability of $G(\cdot)$ in the following proposition. The proof of this proposition as well as all other proofs are provided in the online appendix.

Proposition 1. *In a given auction, the winning probability vector $G_f(b)$ is continuous and differentiable at all b, for all bidders $f \in F$.*

2.1. A Two-Step Structural Estimation Method

For a given bidder, the necessary first-order conditions of the optimization problem (2) are given by the matrix equation

$$c = b + \{[\mathcal{D}_b G(b)]^T\}^{-1} G(b), \tag{3}$$

where $\mathcal{D}_b$ refers to the Jacobian matrix operator with respect to the variable vector b so that the ijth element is $[\mathcal{D}_b G(b)]_{ij} = (\partial/\partial b_j) G_i(b)$.[3] For a given auction, there is one first-order-condition matrix equation per bidder. The standard structural approach assumes that the observed bid vector of each bidder satisfies Equation (3). An important difference between first-price single-unit auctions and CAs is that in the former this first-order condition is necessary and sufficient for optimality, whereas in the latter it is only necessary. However, in principle it is possible to test computationally whether the observed bid vector that satisfies (3) is locally or globally optimal for optimization problem (2). We provide more details in the context of our application.

The first-order conditions (3), evaluated at the observed bid vector in the data, are the basis to point identify that bidder's cost vector, because the right-hand side only depends on the observed bid vector b, the winning probabilities $G(b)$, and their derivatives. Note that Assumption 4 implies that bidders have the correct expectations about the vector of winning probabilities $G(b)$. Hence, these winning probabilities must be consistent with the actual auction play, and therefore can be potentially estimated using bidding data from all bidders. For example, in the first-price single-unit auction analyzed by GPV, the winning probability distribution—which in this case corresponds to the tail distribution of the competitors' minimum bid—and its derivative can be estimated nonparametrically. GPV replace these estimates in the first-order conditions to obtain point estimates of bidders' costs.

[2] An alternative would be to assume that bidders play a *mixed strategy* BNE in the CA; this is guaranteed to exist. However, we believe that formulating a structural model in terms of pure strategies is more transparent, has a clearer interpretation, and yields simpler identification arguments.

[3] The Jacobian $\mathcal{D}_b G(b)$ is a square matrix that can have nonzero off-diagonal elements because bids from the same bidder compete against each other.

In a CA setting, $G_f(\cdot)$ is a vector of probabilities determined by the bid distributions of competitors $\{H_{f'}(\cdot \mid Z)\}_{f'\neq f}$ and the winner determination problem (1), which has no analytical solution. CP use a simulation-based two-step method to estimate $G_f(\cdot)$ and to then use the first-order conditions to obtain point estimates of the bidders' costs. This procedure can be summarized as follows:

Step 1. Use bid data to estimate the distribution of bids, $H_{f'}(\cdot \mid Z)$, for all firms $f' \in F$.

Step 2. To obtain the cost vector of firm f, c_f, estimate via simulation the vector of winning probabilities $G_f(b)|_{b=b_f}$ and its Jacobian matrix $\mathcal{D}_b G_f(b)|_{b=b_f}$ evaluated at the observed bid vector submitted by firm f, b_f. Replace these on Equation (3) to obtain a point estimate of c_f.

In Step 2, winning probabilities are estimated via simulation, where each simulation run $r = 1, \ldots, R$, consists of the following:

- Fix the bid vector by firm f, b_f, and for each competitor $f' \neq f$, independently sample the competitor's bids from the distribution $H_{f'}(\cdot \mid Z)$ estimated in the first step. Let $\{b^r_{f'}\}_{f'\neq f}$ be the bids sampled for each competitor firm in simulation run r.
- Solve the winner determination problem with bid vectors $(b_f, \{b^r_{f'}\}_{f'\neq f})$. Record the package won by firm f, if any, with indicators $\mathbf{1}[win^r_{af}] = 1$ if and only if firm f wins package a in run r.

The winning probabilities are estimated with the empirical frequency of wins over all runs in the simulation, that is, $\hat{G}_{af}(b_f) = 1/R \sum_{r=1}^{R} \mathbf{1}[win^r_{af}]$, where R is an appropriately chosen large number. The Jacobian matrix of $G_f(b_f)$ is computed numerically using a similar simulation technique together with a finite-difference method.[4]

2.2. Identification

It is helpful to understand what patterns in the bid data drive the identification of the cost estimates from Equation (3). This first-order condition equation implies that the bid vector is equal to a cost plus a markup vector, where the markup vector for bidder f, $-\{[\mathcal{D}_b G(b)]^T\}^{-1} G(b)$, depends implicitly on the competitors' bid distributions estimated in the first step of the structural method, $\{H_{f'}\}_{f'\neq f}$, from the bid data. It is therefore useful to analyze how these distributions affect the estimated markups and the cost estimates. For this purpose, we conducted numerical experiments to analyze how a firm's optimal bid vector depends on the distribution of competitors' bids. The online appendix describes the details of these experiments and the rest of this section provides the main insights derived from them.

Specifically, the experiments seek to understand the impact of different quantities that describe the distribution of competitors' bids on strategic markup reductions; we consider the average unit prices, their correlation, and the magnitude of package discounts. We used a multivariate normal distribution to model the unit prices of competitors' bids and analyzed how the optimal bid prices of a focal bidder change as we vary the parameters of this distribution in small-scale instances. Because the costs of the focal firm are kept constant across the experiments, changes in the optimal bids are purely driven by changes in markups.

The results reveal that strategic markup reductions become larger (1) as the competitors' average unit prices increase, because the focal firm becomes more competitive and can more easily exert market power; and (2) as the correlation among single-unit prices becomes more negative; this effect is related to bundling motives for a multiproduct monopolist. The results also suggest that changes in the magnitude and variance of competitors' package discounts (within a reasonable range) do not affect much the markup reductions.

These numerical experiments calculate the optimal bid vector given the firm's costs and the distribution of competitors' bids. In our estimation procedure, we perform the opposite process; we find the costs that rationalize the observed bids using the first-order conditions of the bidders' optimization problem. In this regard, the results in the numerical experiments provide useful insights about identification of costs. Recall that observed package discounts are the sum of cost synergies plus strategic markup reductions. Hence, the results suggest that as the means of the estimated distribution of competitors' bids become larger, and as the correlation among individual prices become smaller, the estimated markup reductions should increase, and therefore, the fraction of the package discounts explained by cost synergies should decrease. In summary, when estimating the distribution of competitors' bids, correctly capturing the correlation and the heterogeneity of unit prices plays an important role in the estimation of the bidders' cost structures.

2.3. Limitations in Large-Scale CAs

CP were able to effectively use the previous approach in auctions of at most three units. However, there are two significant limitations in using the standard approach in large-scale CAs with more units.

First, in large-scale CAs that are typically found in practice (including our empirical application), firms

[4] To calculate the ath row and sth column element of the Jacobian, one can use a central finite-difference method $[\mathcal{D}_b G_f(b)|_{b=b_f}]_{as} = (\partial/\partial b_s) G_{af}(b)|_{b=b_f} \approx (G_{af}(b_f + he_s) - G_{af}(b_f - he_s))/(2h)$, where e_s is the sth canonical vector—the sth component of e_s is its only nonzero element with size one. The step-size h is an appropriately chosen small value (see, e.g., Glynn 1989). The estimations of $G_{af}(b_f + he_s)$ and $G_{af}(b_f - he_s)$ can be obtained via simulation as above.

may submit hundreds or even thousands of bids. In that case, the bid vectors $\{b_f\}_{f\in F}$, and therefore the distributions $\{H_f(\cdot \mid Z)\}_{f\in F}$ that need to be estimated in the first step, are high dimensional. For this reason, parametric restrictions need to be imposed to make the estimation tractable. However, it is important to allow for sufficient flexibility in these restrictions. In their application, CP developed a reasonable parametric model that balances flexibility with feasibility in the estimation. We extend their approach to large-scale CAs. In particular, in §4 we provide more details about a parsimonious, yet flexible parametric description of the distributions of competitors' bids for CAs that involve geographically dispersed and heterogeneous units as well as scale and density discounts. These distributions are then taken as an input for the second step.

Second, there is a limitation in the second step of the CP approach when applied to large-scale CAs: the high dimensionality of the first-order conditions (3). The dimension of this matrix equation is determined by the number of bids, which increases exponentially with the number of units in the CA. As the number of bids submitted by a bidder gets large, the winning probability of each bid is likely to become very small and the simulation errors in estimating these rare-event probabilities become large. Moreover, Equation (3) requires taking derivatives over a large number of variables; the simulation errors in these quantities may be even larger. These problems may not be resolved by simply increasing the length of the simulation runs, because in the course of these runs one needs to repeatedly solve the winner determination problem, which gets increasingly expensive computationally as the scale increases.[5] Hence, computation of $G(b)$ and $\mathcal{D}_b G(b)$ via simulation becomes quickly intractable as the number of units auctioned increases. The difficulties in estimating $G(b)$ make it also unreasonable to assume that bidders would be able to solve (2) optimally.

An important methodological contribution of this paper is to address the second problem—the high dimensionality of the first-order conditions (3). Our approach imposes reasonable restrictions in the structure of the markups that allow us to reduce the dimensionality of the problem. We describe this approach in detail in the next section.

[5] The winner determination problem in a CA is known to be NP-hard. Using state-of-the-art solvers for integer programming, it takes in the order of seconds to solve a single instance of the winner determination problem in our application—we used CPLEX V12.1 called by a C routine and ran on Columbia Business School's shared cluster, where each machine has eight 2.4 GHz CPUs.

3. The Characteristic-Based Markup Approach for Large-Scale CAs

Our model is based on the approach described in §2. As mentioned above, a significant complication of using this model in large-scale CAs is that the dimensionality of the first-order conditions is too large. We develop an approach to reduce the dimensionality of the problem by imposing additional assumptions on the bidders' bidding behavior that have behavioral appeal and make the estimation approach econometrically and computationally feasible in large-scale CAs.

Notice that in the first-order conditions (3), the markup term $-\{[\mathcal{D}_b G(b)]^T\}^{-1} G(b)$ provides the flexibility to the bidder to assign a different and separate markup to each package. Hence, we refer to this model as the *full-dimension* model. In contrast, we propose that the markup of each bid is specified through a reduced set of package characteristics. Specifically, let w_a be a row vector of characteristics describing package a, with dimension $\dim(w_a) = d$ that could be potentially much smaller than A. The markup for package a is given by the linear function $w_a\theta$, where θ is a (column) vector of dimension d specifying the markup components associated with each package characteristic. Instead of choosing the markup for each package, the bidder now chooses θ. Let $W \in \Re^{A\times d}$ be a matrix containing the characteristics of all packages, so that the ath row of W is w_a. The following assumption, kept throughout the paper, formalizes this simplification to the bidders' bidding behavior.

Assumption 5 (Characteristic-Based Markups). *Consider a given bidder in a particular auction. Its bid vector is determined by $b = c + W\theta$, where W is a fixed $(A \times d)$-dimensional matrix of package characteristics and θ is a d-dimensional vector of decision variables chosen by the bidder.*

Under this assumption, the bidder's optimization problem becomes[6]

$$\underset{\theta\in\Re^d}{\text{maximize}}\,(W\theta)^T G(W\theta + c), \tag{4}$$

whose first-order conditions yield

$$[\mathcal{D}_\theta W^T G(W\theta + c)]^T \theta = -W^T G(W\theta + c). \tag{5}$$

Here again the ijth element of the Jacobian matrix above is $[\mathcal{D}_\theta W^T G(W\theta+c)]_{ij} = (\partial/\partial\theta_j)[W^T G(W\theta+c)]_i = (\partial/\partial\theta_j) W_i^T G(W\theta+c)$, where W_i is the ith column of matrix W. Rearranging and replacing terms, we can solve for the decision vector θ as follows:

$$\theta = -\left\{[\mathcal{D}_\theta W^T G(b)]^T\right\}^{-1} W^T G(b). \tag{6}$$

[6] Note that our approach allows the specification of W to vary across bidders.

As in GPV and CP, this first-order-condition equation constitutes the basis of identification in our structural model. Again, note that in each auction there is one first-order-condition matrix equation per bidder and different bidders may have different characteristic matrices W. For each bidder, under Assumption 5, the cost is given by $c = b - W\theta$. Hence, costs are uniquely determined by θ, and moreover, if the matrix $\mathcal{D}_\theta W^T G(b)$ is invertible, Equation (6) uniquely identifies the markup vector θ. Therefore, Equation (6) provides an alternative to (3) to estimate costs.[7] We formalize this discussion with the following assumption that is kept throughout the paper.

Assumption 6 (First-Order Conditions). *The observed bid vector of a given bidder in the auction satisfies the necessary first-order conditions of the characteristic-based markup model given by Equation* (5).

As with Equation (3), the right-hand side of Equation (6) can be estimated purely from observed bidding data. In fact, our approach using Equation (6) closely follows the two-step method described in §2.1. However, the reduced dimensionality of Equation (6) significantly simplifies the computational burden in the second step, making it feasible in large-scale applications.

To see this, note that (6) is similar to (3), with the winning probability vector $G(b)$ and its Jacobian matrix $\mathcal{D}_b G(b)$ replaced by the vector $W^T G(b)$ and its Jacobian matrix $\mathcal{D}_\theta W^T G(b)$, which is now with respect to the markup vector θ. The first simplification is that the derivatives are now taken with respect to $d \ll A$ variables, effectively reducing the dimension of the problem. Second, in the specifications we propose later, we will see that each element of the vector $W^T G(b)$ is a (weighted) sum of winning probabilities over many packages. These aggregate probabilities are larger than the winning probabilities of each individual package, and therefore easier to estimate via simulation. Besides, there are fewer probabilities to be estimated; altogether these make the second step computationally tractable.

One apparent limitation of Assumption 5 is that the markup is additive as oppose to multiplicative to costs, which may be more appropriate in some applications. A multiplicative markup, however, would lead to different first-order conditions from which it is mathematically intractable to identify bidders' costs using bid data. A relatively simple way to make the additive assumption less restrictive is to include package characteristics in W which are related to costs, so that the markup can be scaled based on these cost-related characteristics. This approach is effective when the cost heterogeneity across packages can be captured, at least partially, by a reduced set of known variables. We come back to this point in the sequel.

Note that the characteristic-based markup model is very general and flexible in the specification of markup structures. For example, if we specify the package-characteristic matrix W as the identity matrix, each package has its own markup and we are back to the full-dimension problem (2). On the opposite extreme, one could choose $d = 1$ so that the markups of all packages are determined by a single decision variable; although this specification significantly reduces the dimension of the problem, this may be too restrictive. Between these two extremes there are many possible specifications for W. Different specifications may be chosen depending on the details of the large-scale application at hand. Next, we describe an approach to specify W that is sufficiently flexible to capture strategic markup reductions that arise in package bidding, but that at the same time is parsimonious and maintains computational tractability.

3.1. Specifying Markup Restrictions

Recall from our discussion in the introduction that an important objective of our structural estimation approach is to measure what portion of the observed package discounts can be attributed to cost synergies versus strategic markup reductions. Therefore, in order for the structural estimation to provide meaningful estimates of cost synergies, it is crucial that the markup restrictions are sufficiently flexible to incorporate the main drivers of strategic bidding behavior. Previous literature suggests that scale is likely to be the main driver of these strategic markup reductions (see the references in the introduction, in particular Olivares et al. 2012, for a more detailed discussion). Hence, we focus on developing a specification that allows for markups to vary on the size of the packages, which allows separating what portion of the volume discounts observed in the bid data arises from strategic markup reductions vis-à-vis cost synergies.

Recall that the key idea in our estimation approach is to impose restrictions on the markup structure to reduce the dimensionality of the bidders' problem. A special case of the characteristic-based markup approach is to create a partition of the set of all packages, and allow each group (or set) of the partition to have its own separate markup parameter. This group markup parameter then determines the markup of all the packages in the corresponding group. This approach, referred to as the *group-based markup model*, is defined formally as follows.

Definition. A markup specification follows a *group-based markup model* if each row of the package-characteristic matrix W is composed by zeroes except for one and only one positive component.

[7] In §3.3 and the online appendix, we provide conditions for the invertibility of this matrix.

Consider the following group-based markup model. Let $\{\mathcal{A}_s\}_{s=1}^{S}$ be a partition that covers all possible packages. From this partition, a potential candidate for the package-characteristic matrix $W \in \Re^{A\times S}$ can be generated using indicator variables $W_{as} = \mathbf{1}[\text{package } a \text{ belongs to set } \mathcal{A}_s]$. With this specification, the term $W^T G(b)$ in Equation (6) has the following form:

$$W^T G(b) = \begin{bmatrix} W_1^T G(b) \\ W_2^T G(b) \\ \vdots \\ W_S^T G(b) \end{bmatrix} = \begin{bmatrix} \text{probability of winning any package in } \mathcal{A}_1 \\ \text{probability of winning any package in } \mathcal{A}_2 \\ \vdots \\ \text{probability of winning any package in } \mathcal{A}_S \end{bmatrix}.$$

As seen above, the group-based markup model could significantly reduce the dimensionality of the problem; if $S \ll A$, there are much less probabilities to estimate, as well as derivatives to take in the Jacobian matrix. Moreover, whereas the winning probability of any given package a is typically small and hard to estimate via simulation, the winning probability of a *group* of packages aggregates these individual probabilities over a potentially large set of packages and, therefore, is often much larger. For this reason, we require fewer simulation runs to obtain precise estimates of these aggregate probabilities. All this makes the computation of the right-hand side of the first-order conditions (6) tractable.

A special case of the group-based markup model is when the packages are grouped by their sizes. For some defined measure of package size (e.g., the number of units in the package), let $\mathcal{A}_s$ be the set of all packages of size s. The markup parameter θ_s represents the *common* markup charged to all packages of size s; the bidder chooses S different markups, one for each possible size. This is referred to as the *pure size-based markup model*. Recall that we want to disentangle what portion of the observed bid discounts is explained by markup reductions when bidders submit larger packages. The pure size-based markup model provides the minimum level of flexibility to capture such strategic markup reductions, and therefore, we believe it is a reasonable starting point to impose markup restrictions in our approach.

Now, we seek to understand whether the pure size-based markup model provides a good approximation to the estimates of the full-dimension model. To do so, we provide an analytical comparison of the markups estimated by the full-dimension model with those estimated via the group-based markup approach.

Proposition 2. *Consider a bidder submitting a bid vector b in a CA. Assume that all bids in b have strictly positive probabilities of winning.*

(a) *Suppose the CA has A packages. Let θ_a, $a = 1, \ldots, A$, be the estimated markup for package a by the full-dimension model* (3), *and θ_u be the common markup estimated by the group-based markup model* (6) *when the A packages form a single group, that is, $b_a = c_a + \theta_u$, $a = 1, \ldots, A$. Then, $\theta_u = \sum_{a=1}^{A} \beta_a \theta_a$, for appropriately defined weights $\beta_a \geq 0$, $a = 1, \ldots, K$, that satisfy $\sum_{a=1}^{A} \beta_a = 1$.*

(b) *Suppose the CA has two units. Let $(\theta_1, \theta_2, \theta_{12})$ be the estimated markup vector by the full-dimension model* (3) *and let (θ_u, θ_v) be the estimated markup vector by the pure size-based markup model* (6), *where θ_u is the common markup for single units and θ_v is the markup for the package. Then, $\theta_u = \beta\theta_1 + (1-\beta)\theta_2$ and $\theta_v = \theta_{12} + \gamma(\theta_1 - \theta_2)$, for appropriately defined constants $\beta \geq 0$ and γ.*

We provide the proof of the above proposition in the online appendix, as well as the detailed expressions for the constants β's and γ. The results from Proposition 2 provide important insights regarding the implications on the group-based markup estimates. First, from the result in part (b), we observe that grouping the units affect the estimated markup of the package, where the impact depends on the coefficient γ and the difference of the individual unit markups. If the unit markups are very close to each other, the effect of grouping on the package markup will be negligible. Moreover, note that if γ is small, the effect of having a common markup for the units has a negligible effect on the estimated markup for the package. In fact, extensive numerical experiments have shown that in our application, grouping a set of packages so that they share a common markup merely affects the markups of other packages that are not in that particular group.[8]

The previous discussion together with part (a) in the proposition can be summarized as follows. Consider a situation in which the packages in the set $\mathcal{A}_s$ are grouped together, and let θ_s be the common markup estimated by the group-based markup model. Let θ_a be the individual markup for package $a \in \mathcal{A}_s$ estimated by the full-dimension model. Then, the previous discussion basically suggests that

$$\theta_s \approx \sum_{a\in\mathcal{A}_s} \beta_a \theta_a,$$

where $\beta_a \geq 0$, $\forall\, a \in \mathcal{A}_s$, are appropriately defined weights that satisfy $\sum_{a\in\mathcal{A}_s} \beta_a = 1$. The result is useful

[8] We validate these observations empirically in §6.1 using real data from small-scale CAs.

because it suggests that the estimated common markup is a convex combination of the individual markups we would obtain from the full-dimension model (if we were able to estimate them). Moreover,

$$|\theta_a - \theta_s| \approx \left|\theta_a - \sum_{a' \in \mathcal{A}_s} \beta_{a'} \theta_{a'}\right| \leq \sum_{a' \in \mathcal{A}_s,\, a' \neq a} \beta_{a'} |\theta_a - \theta_{a'}|, \quad \forall\, a \in \mathcal{A}_s.$$

Therefore, the estimated common markup would be a good approximation to the individual markup estimates from the full-dimension model if the latter markups are close to each other. Of course, checking this condition is computationally intractable, because we would need to solve the full-dimension model. The next section describes a computationally tractable heuristic that aims at providing more flexibility in the markup restrictions without increasing much the computational burden of the method.

3.2. A Refinement of the Size-Based Markup Model

As suggested above, the main issue with the size-based markup model would be whether or not the packages in the same size group are significantly heterogeneous. For example, if one package has a significantly different markup to the rest of the group in the full-dimension model (if we were able to estimate it), it ideally should not be a part of the group.

Recall that one of the difficulties in the full-dimension model arises in the computation of a large number of small winning probabilities via simulation. However, for a given firm, there still may be a small number of packages with reasonably large winning probabilities that can actually be computed with precision—we refer to these as "special packages." It is then possible to assign and estimate a separate markup for these special packages, without forcing them into a group and therefore alleviating the potential biases previously discussed. Given their high winning probabilities, special packages are also more likely to be part of the winning CA allocation, so it is useful to obtain more precise estimates for their markups. Finally, in §6.1 we provide empirical evidence in the context of our application that high winning probability packages tend to have larger estimated per-volume markups in the full-dimension model relative to the rest in their corresponding size groups. Hence, removing them from the groups and estimating a separate markup for each of them is likely to reduce the bias associated with grouping in a significant way.

Another extreme alternative would be to estimate the model with special packages only, ignoring the rest of the packages. Although the cost information of bids with small winning probabilities may be less important for the estimation of the performance measures (since they are less likely to be part of the winning CA allocation), they cannot be entirely eliminated in the estimation procedure. One reason is that these packages may have significant winning probability in aggregate, and therefore ignoring them in the first-order conditions (6) can result in an inaccurate estimation for the costs of the large winning probability bids. In fact, we have estimated models with and without the packages with small winning probabilities and found that the markup estimates of the special packages changed substantially.

In addition, in some applications (including the one analyzed in this work) the units of the same group can be heterogeneous and this could lead to differences in markups, even after separating the special packages. In our application, units differ in their volume and so packages with the same number but different composition of units could have different markups. To account for this heterogeneity, let v_i be the volume of unit i and define $v_a = \sum_{i \in a} v_i$ as the total volume of the package. The package-characteristic matrix W can be specified as $W_{as} = v_a \cdot \mathbf{1}[\text{package } a \text{ has } s \text{ units}]$, which is also in the class of group-based markups. With this specification, the packages in the same size group will share the same markup parameter θ_s, which is the per unit of volume markup, so that the markup of package a is $v_a \theta_s$. Also, the sth element of the vector $W^T G$ is equal to the expected volume of winning packages of size s. Overall, this specification makes the additive nature of Assumption 5 less restrictive.

Based on the previous insights, we propose the following heuristic to build the package-characteristic matrix W for a given firm:

1. Group packages according to their sizes and let $W_{as} = v_a \cdot \mathbf{1}[\text{package } a \text{ has } s \text{ units}]$, so that initially all packages with the same number of units share the same markup per unit of volume.

2. Run a simulation to obtain rough estimates of the winning probabilities of each package; this simulation is quicker to run than solving for the first-order conditions. For each size group, identify bids that have high winning probabilities relative to the rest. Each of these packages is associated with a separate individual markup parameter.

3. For each size, further divide the rest of the bids into two groups: medium and low winning probability groups. This step is motivated by the observation discussed in §6.1 that winning probability is related to the magnitude of markups. In that section, we further justify this step with empirical evidence in the context of our application.

Through this heuristic procedure, we construct the corresponding package-characteristic matrix W for each firm allowing for separate markups for each of the specified groups (including the groups with a single special package). We refer to this approach to define the package-characteristic matrix as the *extended size-based markup model*, which is a particular case of a group-based markup model.[9] For each firm, we use this specification within a two-step method similar to the one described in §2.1. In the first step, we parametrically estimate the distributions of competitors' bids. This procedure requires a separate treatment and is described in detail in §4. In the second step, we use the specification of W given by the previously described heuristic in the first-order conditions (6) to obtain a point estimate of θ, and hence of c.

Our heuristic based on the extended size-based markup model aims to improve the approximation to the full-dimension model starting from the pure size-based markup model. However, it is important to provide some empirical validation of this claim. For this purpose, we collected data from two exceptionally small CAs in our application. The full-dimension approach was feasible to implement in these smaller auctions and was compared with the results provided by the extended size-based model using our heuristic method described above. Notably, the results of this analysis presented in §6.1, suggest that the markups estimated with the two approaches are very similar, providing support for our method.[10]

3.3. Further Requirements on the Package-Characteristics Matrix

We finish this section by discussing issues related to identification that are important for the specification of W. In particular, we provide conditions for which $\mathscr{D}_\theta W^T G(b)$ is invertible in Equation (6), and therefore, the first-order conditions uniquely identify the markup vector θ, and hence the costs.

In some applications, including the one analyzed in this paper, bidders may not submit bids on all packages.[11] This case can still be handled with our proposed approach by treating the missing packages as bids with very high prices that have no chances of winning. We refer to the bids that never win as *irrelevant bids*. In addition, submitted bids with zero probability of winning are also considered irrelevant. In contrast, a *relevant bid* has a strictly positive probability of winning.

CP show that in the full-dimension model, irrelevant bids do not play a role in the first-order conditions, and one can identify the markups for relevant bids after eliminating irrelevant bids from the estimation. This result extends to the group-based markup model, as long as each group has at least one relevant bid. Theorem 1 in the online appendix provides necessary and sufficient conditions for the invertibility of the Jacobian matrix $\mathscr{D}_\theta W^T G(b)$. A practical implication of the theorem is that when implementing the heuristic described in §3.2 one needs to make sure that each group of packages must include at least one relevant bid. After we imposed this, we were always able to invert the Jacobian matrix computationally. Another implication of the theorem and the related discussion in the online appendix is that the proposed method can only identify the cost structure of packages associated with relevant bids, because irrelevant bids provide no information to the first-order conditions. We come back to this point in §6 in the context of our application.

Finally, an important assumption needed for our approach is that bidders can win at most one package. This is a frequent requirement in many real-world CAs, especially in settings with rich and expressive package bidding. Without this requirement, it may not be possible to point identify costs. For example, consider a CA with two units and suppose a bidder only submits bids for the individual units. Suppose the bidder has a positive chance of winning both individual bids simultaneously, which is equivalent to winning the two-unit package. Then, we have three unknowns to estimate (the cost for each individual unit and the cost for the package), but only two equations (the two first-order conditions with respect to the individual bid prices).

4. Estimating the Distribution of Competitors' Bids

Let us first recapitulate the two-step approach introduced in §2. In the first step, we need to estimate the distributions of competitors' bids, $\{H_f(\cdot \mid Z)\}_{f\in F}$, which are then used in the simulation-based routine in the second step to sample competitors' bids and estimate the terms in the first-order conditions given by Equation (6). Section 3 addressed the complexity introduced in the second step due to the large-scale nature of the auction. In particular, we simplified

[9] We try to have as many markup parameters as possible to the extent that computational tractability is maintained. In our actual estimation, we use the threshold probability of 10^{-3} to identify the high probability special package bids. Packages with winning probabilities above 10^{-4} are categorized in the medium probability groups, and the rest in the low probability groups. With this grouping procedure, a typical firm has a markup vector with dimension $d = 20$.

[10] We note that in the related context of multiproduct monopolist pricing, Chu et al. (2011) provide computational and empirical evidence of the effectiveness of size-based pricing in some settings. They also show examples where this restricted pricing strategy is used in practice.

[11] In fact, in our empirical application, firms do not place bids on all possible combinations because of two reasons: (1) firms have limits on the maximum number of units that can be included in a package (these limits depend on the firm's financial capacity); and (2) the number of possible combinations is too large.

the first-order conditions of the bidder's problem by imposing some structure in their markups. It was important that the structure was flexible enough to allow for strategic markup reductions.

The complication in the first step is that in large-scale CAs, firms may submit hundreds or even thousands of bids. Therefore, the bid vectors $\{b_f\}_{f\in F}$ are high dimensional precluding the use of a nonparametric approach like GPV to estimate the distribution of competitors' bids; CP faced a similar challenge even for a three-unit CA. This section describes a parametric approach to model the bid distribution that can be used in CAs that involve geographically dispersed and heterogeneous units that are subject to discounts due to scale and density, like in our application.

It is important to emphasize that the simplifications in the two steps have different objectives. In the first step, the objective is to introduce a parametric model that fits the competitive bidding landscape data well. In the second step, the objective is to simplify the bidders' decision space in the first-order conditions. We note that the parametric model of the *bid* data in the first step will be more flexible than the model for *markups* in the second step, because it will allow for scale and density discounts both of which could be observed in the data. On the other hand, our extended size-based markup model explicitly considers strategic discounts associated with scale only. The reason is that, as mentioned in §3.1, economic theory suggests that scale (and not density) is likely to be the main driver of the strategic markup reductions we are trying to identify. In §6.1 we provide some empirical evidence of this claim in the context of our application.

The parametric approach we follow has an important difference with CP in that in our estimation method the identification of the distribution of competitors' bids is based on variation across package bids and firms in a *single auction*, and hence exploits the large number of package bids, which is a key characteristic of large-scale CAs. In the standard structural approach to auctions (including CP and GPV), the estimation of the bid distribution uses variation in a *cross-section of auctions*, implicitly assuming that the same equilibrium is being played across these auctions. Hence, our identification strategy can be more robust when there is unobserved heterogeneity across auctions—changes in the auction characteristics and firm characteristics from auction to auction that are observed by bidders but unobserved by the econometrician (for a more detailed discussion on this issue, see Krasnokutskaya 2011).

Imposing parametric restrictions to the multivariate bid distribution needs to balance flexibility with estimation feasibility. There are three key aspects typical in applications of CAs that are important to account for: (i) heterogeneity among units; (ii) the correlation structure among the bids from the same bidder; and (iii) package discounts. We discuss each of these three in what follows.

First, in many CAs, the bid prices are heterogeneous among units and among firms. In applications that involve logistics and transportation across dispersed geographic units (as the one we study), heterogeneity among units arises primarily from the costs of serving different territories. For example, units located in isolated rural areas tend to be more expensive than units in urban areas. There is also heterogeneity across firms: some firms may have national presence, are vertically integrated, and may have well-functioning and efficient supply chains; other firms may be more rustic local firms.

Second, package bids of the same bidder may be correlated. In CAs, there are two main factors that can generate correlation between bids. First, a bidder that has a high cost in a given unit is likely to submit higher prices for all packages containing that unit. Second, if there are local advantages, a supplier charging a low price for a unit may also charge lower prices for nearby units. Hence, the unit composition of the package bids together with the spatial distribution of the territorial units provides a natural way to parameterize the covariance structure among package bids. As described in §2.2, the correlation structure of the competitors' bids has direct implications on the incentives to engage in strategic markup reductions, so it is important to allow for a flexible covariance structure that incorporates these effects.

Third, CAs exhibit package discounts in the bids; the price per unit may decrease as the size of the package increases. In applications where economies of density matter, the geographic location can be another factor that determines the magnitude of the discounts; for example, combining two units located nearby could lead to larger discounts (relative to a package with two distant units).

Accordingly, we develop the following econometric model for package bids that captures heterogeneity among units, correlation, and discounts. In particular, from the perspective of all other firms, firm f's bids are specified by the following parametric model:[12]

$$b_{af} = -g^{\text{scale}}(v_a, \beta^{\text{scale}}_{k(f)}) - \sum_{c\in \text{Cl}(a)} g^{\text{density}}(v_c, \beta^{\text{density}}_{k(f)}) \cdot \frac{v_c}{v_a} + \sum_{i\in a} \tilde{\delta}_{if} \frac{v_i}{v_a} + \tilde{\varepsilon}_{af}. \tag{7}$$

As defined earlier, v_i denotes the volume of unit i and $v_a = \sum_{i\in a} v_i$ is the total volume of package a.

[12] The structure in Equation (7) that separates individual prices with discounts is motivated by Olivares et al. (2012).

With some abuse of notation, the dependent variable, b_{af}, denotes the price per unit of volume submitted by firm f for package a; that is, the actual bid price divided by the total volume of the package, v_a. The four terms in the right-hand side of Equation (7) capture (i) the effect of discounts due to size or scale (g^{scale}); (ii) the effect of discounts due to density (g^{density}); (iii) the effect of the specific units contained in the package (the sum over units i in package a), where $\tilde{\delta}_{if}$ can be viewed as an *average implicit price* that bidder f is charging for unit i among all the packages submitted, net of any scale and density discounts; and (iv) a Gaussian error term $\tilde{\varepsilon}_{af}$ capturing other factors affecting the bid price. It is important to emphasize that the discount functions g^{scale} and g^{density} should not be interpreted directly as cost synergies because part of the discounts could also arise from strategic behavior.

This parametric specification also assumes that the bids across bidders are independent and that the bid distribution of a bidder depends only on its own characteristics, $H_f(\cdot \mid Z) = H_f(\cdot \mid Z_f)$.[13] In addition, to avoid making strong assumptions on how firms choose which combinations to submit, we use the same package composition observed in the data. That is, when generating competitors' bids, we fix the packages on which the bids were actually submitted by a particular bidder and simulate new prices for these packages.

The competitors' bid distribution captures the relevant uncertainty faced by a bidder because of asymmetric information in the auction game. Hence, it is important to distinguish which elements of Equation (7) are known by all other firms at the time of bidding and which are private information to firm f submitting the bid vector. We use tilde (e.g., $\tilde{\delta}_{if}$) to denote factors that are private information to firm f and therefore treated as random parameters from the perspective of all other bidders. As a consequence, the bid distribution $H_f(\cdot \mid Z_f)$ is characterized by the deterministic parameters $\{\beta^{\text{scale}}_{k(f)}, \beta^{\text{density}}_{k(f)}\}_{f\in F}$ (to be defined shortly) and the *distribution* of the random parameters $\{\tilde{\delta}_{if}\}_{i\in U,\, f\in F}$ and $\{\tilde{\varepsilon}_{af}\}_{a\in\mathcal{A},\, f\in F}$. This distinction between deterministic and random parameters in Equation (7) is important for simulating winning probabilities. Next, we provide more details on how these different components are specified and estimated.

4.1. Model Specification and Estimation Method

First, consider the terms capturing scale and density discounts, $(\beta^{\text{scale}}_{k(f)}, \beta^{\text{density}}_{k(f)})$. The model allows for some observed heterogeneity of these discounts across firms, with $k(f)$ indicating the type of firm f. For example, firms could be categorized based on their business size, because larger firms may operate at a different cost scale and therefore their synergies could be different. Moreover, larger firms may also be able to bid on larger packages, so their markup reductions could also be different. We assume that the heterogeneity in the discount curves across firms is considered common knowledge and that all the uncertainty associated with the magnitude of the discounts is provided by the error terms $\tilde{\varepsilon}_{af}$.[14]

To measure scale discounts, g^{scale} is specified as a step function of the package volume v_a. Because density discounts depend on the proximity of the units in the package, g^{density} depends on the volume of *clusters* of units in a package, where a cluster is a subset of the units in package a that are located in close proximity. In Equation (7), $\text{Cl}(a)$ denotes the set of clusters in the package and c indicates a given cluster in this set, with volume v_c. This approach follows directly from Olivares et al. (2012), and further details on a specific way of computing clusters used in our application is described in the appendix of that paper.

Consider now the term $\sum_{i\in a}(v_i/v_a)\tilde{\delta}_{if}$, a weighted average of firm-unit specific random parameters that capture the effects of the individual units contained in package a. The $\tilde{\delta}_{if}$'s are average implicit prices that bidder f charges for each unit among all the packages submitted, net of any discounts. These implicit prices capture heterogeneity in the unit characteristics (e.g., urban versus rural units) and local advantages of a firm in that unit, among other factors. Part of the heterogeneity of these implicit prices is considered to be private information. Accordingly, we let the vector of average implicit prices $\tilde{\delta}_f$ follow a multivariate normal distribution with mean and covariance matrix (μ, Σ). More specifically, let

$$\tilde{\delta}_{if} = \bar{\delta}_i + \beta^Z Z_{if} + \tilde{\psi}_{r(i),\,f} + \tilde{\nu}_{if}, \tag{8}$$

so that $\mu_i = E(\tilde{\delta}_{if}) = \bar{\delta}_i + \beta^Z Z_{if}$ is specified by a unit fixed effect and the firm characteristics Z_{if}. Firm characteristics depend on the specific application, but may include an indicator on whether the firm was awarded the unit in the previous auction and other covariates that capture local advantages of the firm. The error terms $(\tilde{\psi}_{r(i),\,f}, \tilde{\nu}_{if})$ impose restrictions on the covariance matrix Σ based on the spatial location of units. Let $\mathcal{R}$ be a set of regions that cover all the units in U and $r(i)$ denote the region that contains

[13] Assuming that the bid distribution of a firm depends only on its own characteristics is not restrictive when the distribution is estimated separately for each auction, because the characteristics of the competitors are held fixed within the auction.

[14] In §2.2, through small experiments we also found that modeling these discount parameters as random variables does not affect the cost estimates by much, given the magnitude of discounts observed in the data in our application.

unit i; the number of regions, R, is smaller than the number of units. Each firm is associated with a realization of the random vector $\tilde{\psi}_f = (\tilde{\psi}_{1f}, \ldots, \tilde{\psi}_{Rf})$ from a multivariate normal distribution with zero mean and covariance matrix Ω. The error term $\tilde{\nu}_{if}$ follows an independent, heteroscedastic, zero-mean normal distribution with variance σ_i^2.

Under the specification (8), the covariance structure of any two average implicit prices $\tilde{\delta}_{if}$ and $\tilde{\delta}_{jf}$ is given by $\text{Cov}(\tilde{\delta}_{if}, \tilde{\delta}_{jf}) = \Omega_{r(i),\, r(j)} + \sigma_i \sigma_j \mathbf{1}[i = j]$. Thus, under this model, two unit prices will be more positively correlated if the regional effects of the corresponding regions are more positively correlated. Note that this specification imposes positive correlation among unit prices in the same region; this restriction can be validated with data from the specific application. The model is flexible in allowing positive or negative correlation among units in different regions. Because R may be much smaller than the number of units, this specification provides a substantial dimensionality reduction over the fully flexible covariance matrix Σ.

In summary, the competitor's bid distribution $H_f(\cdot \mid Z_f)$ is a mixture defined by Equations (7) and (8), $\tilde{\psi}_f \sim \text{MVN}(0, \Omega)$, $\tilde{\nu}_{if} \sim \text{N}(0, \sigma_i)$, and the error $\tilde{\varepsilon}_{af}$, which is assumed to have a zero-mean normal distribution with variance dependent on the package size, $\sigma^2_{\varepsilon,\, |a|}$. We seek to estimate the vector parameters β^{scale}, β^{density}, $\bar{\delta} = (\bar{\delta}_1, \ldots, \bar{\delta}_N)$, β^Z, $\sigma^2 = (\sigma_1^2, \ldots, \sigma_N^2)$, the covariance matrix Ω, and $\{\sigma^2_{\varepsilon,\, |a|}\}$ for different package sizes. The following two-step method is used to estimate these parameters:

- First step: Estimate (7) via a generalized least squares (GLS) regression to obtain estimates of β^{scale}, β^{density}, $\{\sigma^2_{\varepsilon,\, |a|}\}$, and point estimates of the realizations of the average implicit prices $\tilde{\delta}_{if}$'s.
- Second step: Replace the estimated $\tilde{\delta}_{if}$'s into Equation (8) and estimate the parameters characterizing its multivariate-normal distribution through maximum likelihood.

Identification of the parameters is based on variation across units and firms within a single auction. More specifically, the estimation of the scale and density discounts uses variation across different combinations submitted by the same firm over the same set of units. Given consistent estimates of the realized implicit average unit prices $\tilde{\delta}_{if}$, the second step provides consistent estimates of $\{\bar{\delta}_i, \sigma_i\}_{i \in U}$, β^Z, and Ω as long as Z_{if} is orthogonal to the error components $\tilde{\psi}_{r(i),\, f}$ and $\tilde{\nu}_{if}$. The consistency of our two-step method is a special case of the two-step M-estimators described in Wooldridge (2002).

The next section describes an application of the structural model, and provides further specification details on how to model the bid distribution and the estimates obtained.

5. Application: The Chilean Auction for School Meals

The application we study in this paper is the Chilean auction for school meals. In this section, we start by providing a detailed description of the auction and the data collected. Then, we justify the assumptions of the structural model in the context of our application, and we present the empirical results describing the distribution of competitors' bids.

Junta Nacional de Auxilio Escolar y Becas (JUNAEB) is a government agency in Chile that provides breakfast and lunch for 2.5 million children daily in primary and secondary public schools during the school year. This is one of the largest and most important social programs run by the Chilean government. In fact, in a developing country where about 14% of children under the age of 18 live below the poverty line, many students depend on these free meals as a key source of nutrition.

Since 1999, JUNAEB assigns its school meal service contracts through a single-round, sealed-bid, first-price CA, that was fully implemented for the first time that year. The CA has been used every year since its inception awarding more than US$3 billion of contracts, being one of the largest state auctions in Chile (in recent years, each auction awards contracts for about half a billion dollars). The auction process begins with the registration of potential suppliers followed by an evaluation conducted by the agency, which considers managerial, technical, and financial performance metrics. Some companies may be excluded from the auction if they do not pass this evaluation. Meal plans and service quality are standardized, so that qualified suppliers compete on price.

For the purpose of the auction, Chile is divided into approximately 100 school districts or territorial units (TUs) in 13 geographic regions. Each year, JUNAEB holds an auction for one-third of the country (around 30–35 TUs), awarding three-year contracts. Typically about 20 firms participate in each auction and they are allowed to submit package bids that cover any combination of TUs and specify the prices to serve them. The maximum number of TUs that a firm is allowed to include in any given package (ranging from one to eight TUs) depends on the firm's financial evaluation. Vendors can submit many bids and each package bid is either fully accepted or rejected (i.e., the mechanism does not allocate a fraction of a bid); most firms submit hundreds or even thousands of bids. Two potential sources of cost synergies motivate the use of CAs in this context: (i) economies of scale, generated by volume discounts in the input purchases; and (ii) economies of density that arise from common logistics infrastructure used to supply nearby units.

Firms submit their package bids in a single-round sealed-bid format. Contract winners are paid the amount of their winning bids and are responsible for managing the entire supply chain associated with all meal services in the awarded TUs. This includes from sourcing food inputs, delivering them to the schools, cooking the meals, and serving them to the children. The winning allocation is chosen by selecting the combination of bids that cover all the TUs in the auction at a minimum procurement cost for the government. This winner determination problem is formulated as an integer program (IP) that incorporates other considerations and side constraints (see the online appendix). These constraints impose (i) that no firm can be awarded more than 16% of the country at any point in time (considering all auctions in the past three years); (ii) that firms cannot win more than the number of TUs established by their financial evaluation; (iii) a minimum and maximum number of firms that can operate in any given geographic region; (iv) a minimum number of winning firms in every auction (usually around 10).

Data were collected and processed for all auctions between 2002 and 2005. The data set contains all bids placed by all firms in each auction, the identity and characteristics of participating firms in each auction, and detailed information on the auction parameters, including all the parameters used to determine the side constraints of the winner determination problem. TU data includes its annual demand (number of meals to be served), referred to as the *volume* of the TU, as well as the geographic location of its schools. We also know the set of winning bids in each auction and therefore, at every point in time, we know the identity of the firms serving each TU. Additional details of the data can be found in Olivares et al. (2012).

We apply our method to the large-scale CA of 2003. In 2002, the auction faced some regularity issues, and a second subsequent auction was used to award the contracts. Hence, we conservatively decided to exclude this year from our analysis. In 2004, the government introduced an electronic bidding system to the auction process that resulted in a huge increase in the number of submitted bids. On average, firms placed four times as many bids as they did in 2003, imposing an onerous amount of computation time in the estimation. However, the estimation was more manageable for the 2005 auction as the number of units auctioned and the participating firms were smaller, so we estimated that year to cross-validate the results. Table 1 provides summary statistics of the 2003 auction.

Additionally, we also collected data from two exceptionally smaller-scale auctions that were run between 1999 and 2005. These auctions were used to replace contracts from a few firms that had some irregularities. The auctions had eight and six units, respectively, and about 13 firms participated. Given their smaller scales, these auctions can be estimated with the full-dimension model described in §2. We used them to compare results between the full-dimension and the extended size-based markup models, thereby providing validation of the methods developed in this work. Section 6.1 reports this comparison.

5.1. Discussion on the Assumptions of the Structural Model

In this section, we discuss how the assumptions of the structural model fit into this application. First, Assumption 1 ensures that the auction allocates at most one package per firm. Although this restriction is not explicitly imposed in our empirical application, firms actually win at most one package in practice. In fact, among 41 winning firms between 2002 and 2005, only in one occasion (in 2002) a firm won more than one package (we provide the detailed data in the

Table 1 Summary Statistics for the 2003 Auction

Bidder characteristics					
Business size	Small (1–2)	Medium (3–4)	Medium–Large (5–6)	Large (7–8)	Total
No. of bidders	5	5	2	8	20
Average no. of bids per bidder	308	817	2,540	3,718	2,022

Unit (TU) characteristics						
Region	4	5	9	12	13	Total
No. of TUs	5	10	9	1	7	32
Average volume	2.50 (0.62)	2.32 (0.47)	2.62 (0.67)	2.03 (–)	2.78 (0.70)	2.52 (0.60)
No. of bidding firms	17	19	18	14	19	20

Notes. In the top panel, business size is measured by the number of TUs allowed to win, which is specified next to the business size type; e.g., medium firms are those that can win up to three or four TUs. The bottom panel displays a summary for each of the five geographical regions that were part of the auction. "No. of TUs" refers to the number of territorial units in each region. The volume of the TUs are shown in million meals per year and standard deviations are shown in parentheses. "No. of bidding firms" refers to the number of firms that submitted package bids containing at least one unit in the region.

online appendix).[15] Hence, imposing this assumption is reasonable in this setting.

Assumption 2 imposes independent private costs, which seems adequate in this application. Roughly, 75% of the cost structure of firms is associated with food inputs and labor. A significant amount of these costs are common to all firms. However, this common part is not subject to uncertainty; it is determined by food prices and wages that are common knowledge to all parties involved at the time of the auction (wages of the cooks in this industry are actually regulated by the government). There could still be some uncertainty about future food prices because of the three year extension of the contracts. However, if prices change too much, there are rules in the auction that allow *all* firms to adjust their bids accordingly, dramatically reducing this risk (these rules are based on variation of food price indexes). Therefore, we believe the cost uncertainty is basically driven by firm specific differences in (1) logistics and management abilities (constitute the other 25% of the costs); and (2) idiosyncratic cost advantages related to food inputs, such as better contracting terms with providers. We think (1) and (2) are well captured by an independent private cost model.

Finally, Assumption 4 imposes that firms have the correct expectations regarding the vector of winning probabilities given their bids and competitors' strategies. As previously discussed, this is similar to assuming equilibrium play as is usually done in the structural estimation literature. Although this may generally be a strong rationality assumption given the complexity of the auction, we believe that in our application it may be less so. First, these auctions are repeatedly run every year and all past bidding history is publicly available (including winning and nonwinning bids). In this regard, we exclude auctions where the units were awarded for the first time (years 1999–2001), because bidders had less experience and history to rely on, and were less sophisticated, so that our structural model assumptions may be harder to justify.[16] Also, note that under our characteristic-based markup model firms do not need to estimate the winning probabilities of each individual package, but instead aggregate probabilities over several packages, which can be more manageable.

Moreover, we know from anecdotal evidence that firms in our application are quite sophisticated when bidding. In fact, because stakes are so high, firms invest important amounts of money in business intelligence. Using the historical information together with current market intelligence, firms indeed try to estimate the competitive landscape they will face. A personal interview with a former CEO of one of the supplying companies (that also consulted for other firms) provided more details about the bidding process. He began by creating a large spreadsheet with the packages he was interested in and calculated detailed cost estimates for these. Then, he would choose a markup for the different packages. Historical bid prices were used to decide the average margins to be charged; the margins became smaller over time as the market became more competitive (e.g., when larger catering companies entered the market). In addition, he would typically adjust the markup depending on the number of units in the package, asking for a lower per-meal markup for larger packages. Finally, based on historical bid data, he would also adjust markups for a few packages in which he was "more competitive."[17] This provides further support for Assumption 4 and is also consistent with the extended size-based markup model.

5.2. Estimates of the Bid Distribution Parameters

This section describes the estimates of the distribution of the competitors' bids, based on the model presented in §4. We provide the results for the 2003 auction. Results for the 2005 auction have similar pattern and magnitude, and are omitted for brevity.

The school meals auction exhibits significant differences in discounts across the largest firms and the rest, so we categorize the bidders into two types, $k(f) \in \{L, O\}$ (for large and other), to estimate discounts.[18] On the covariates Z_{if} (see Equation (8)) we include an indicator on whether the firm won the unit in the previous auction (other covariates were also tested but they did not exhibit explanatory power).

Table 2 reports estimates of β^{scale} and β^{density} from the first-step regression (Equation (7)). The scale and density per-meal discount curves, $g^{\text{scale}}(v_a, \beta^{\text{scale}}_{k(f)})$ and $g^{\text{density}}(v_c, \beta^{\text{density}}_{k(f)})$, are specified as step functions with intervals of three million meals per year in the package and cluster volumes v_a and v_c, respectively. Each number indicates the average discount in per-meal price when units are combined to form a package that

[15] Moreover, the government closely monitors the firms that participate in the auction and keeps track of strict records regarding firms' ownerships. Firms that are divided are actually treated as a single firm in the entire auction process. Hence, the government can prohibit firms to win multiple packages through different entities.

[16] On a related point, note that we assume firms maximize expected profits in the current auction without incorporating the impact on future auctions. This assumption is fairly standard in the literature on structural estimation of auctions, and we believe it captures the first-order objective of firms in this market. Moreover, given that our model is already very complex, adding dynamics would make it even more challenging.

[17] The interviewee asked for confidentiality of his identity; we are grateful to this anonymous contributor for the insights provided.

[18] Based on the financial evaluation and business capability, each firm has a maximum number of TUs that it is allowed to win in a given auction. We use this to measure the size of a firm.

Table 2 Results from the First-Step Regression (Equation (7)) for the 2003 Auction

Large firms			Other firms		
Volume	Scale	Density	Volume	Scale	Density
[3, 6]	8.33 (1.30)	6.46 (0.51)	[3, 6]	8.50 (0.62)	1.82 (0.14)
[6, 9]	15.21 (1.33)	7.81 (0.53)	[6, 9]	11.86 (0.64)	3.31 (0.19)
[9, 12]	17.82 (1.31)	8.10 (0.55)	[9, 12]	13.50 (0.65)	3.92 (0.24)
[12, 15]	19.10 (1.30)	8.57 (0.56)	[12, 15]	13.44 (0.67)	5.69 (0.28)
[15, 18]	20.76 (1.29)	9.13 (0.57)	[15, 18]	12.42 (0.69)	6.96 (0.36)
[18, 21]	22.78 (1.30)	11.27 (0.65)	[18, 21]	10.90 (0.72)	
[21, 24]	24.38 (1.30)				
[24, 27]	24.95 (1.35)				

Notes. Robust standard errors are shown in parentheses. Volume is measured in million meals per year, and discounts in Ch$.

belongs to the corresponding volume level. For example, when units are combined into package a with volume $v_a \in [18, 21]$, then on average, a large firm submits a bid that is Ch$22.78 cheaper per meal than the weighted average bid price of those individual units in the package. If all these units are located nearby and form a cluster, there is an additional discount of Ch$11.27 on average for a large firm. The results show that large firms were able to provide higher discounts that can be up to 8.5% of the average bid price (the average bid price in the 2003 auction is Ch$423). All the coefficients are estimated with precision and are different from zero with statistical significance (0.01% significance level). The R-square of the regression corresponding to Equation (7) is 0.98 (with $\tilde{\delta}_{if}$'s as fixed effects), which provides some support that the parametric model adopted provides a reasonable approximation to the bid data generating process.

The second-step estimation (Equation (8)) provides estimates for the distribution of the average implicit prices $\tilde{\delta}_{if}$'s, characterized by $\{\bar{\delta}_i, \sigma_i\}_{i\in U}$, the covariance matrix Ω and β^Z, the coefficients of the firm characteristics. Because of space limitations, we do not report the estimates of the $\bar{\delta}_i$ parameters, but these were estimated with precision—on average, the standard errors are 1.2% of the point estimates. The estimated coefficient for β^Z is -5.986 with a p-value of 0.012, suggesting that on average a firm that was awarded a particular unit in the previous auction submits around 1.5% cheaper bids on packages that contain the unit.

The correlations between the region effects $\psi_{r(i),f}$, calculated based on estimates of the covariance matrix Ω, are provided in the online appendix. These estimates imply a significant positive correlation among units: on average, the correlation between the implicit prices of two units in the same region is 0.68, and 0.45 for units located in different regions. All the standard errors of the maximum likelihood estimates of Equation (8) are computed via a parametric bootstrapping procedure.

We also tested some of the parametric assumptions in our model. First, our approach assumes that the implicit prices δ_{if} follow a normal distribution. For all the units, a Shapiro–Wilks test cannot reject this assumption at 5% significance level (p-values are in the range of 0.073 to 0.92). To test the restrictions on the covariance structure of the implicit prices $\{\delta_{if}\}_{i\in U}$ imposed by Equation (8), we compared this model against a more general model where the full covariance matrix Σ is unrestricted. A likelihood ratio test cannot reject that the two models are equivalent (p-value > 0.6 for both 2003 and 2005 auctions). These tests confirm that our parametric assumptions make the estimation tractable, while being reasonably flexible.

6. Costs, Markups, and Performance Estimation Results

We estimate markups and costs for the school meals auction application using the two-step approach described in §§2 and 3. We use the competitors' bid distribution estimated in §5.2 in the first step, and we apply the heuristic based on the extended size-based markup model described in §3.2 in the second step. Before showing those results, to validate our estimation approach, we compare the estimation via the full-dimension and the extended size-based markup models using the two small CAs.

6.1. Small Auctions Estimation

In this section, we estimate the markups and costs with the full-dimension model and the heuristic based on the extended size-based model in the small-scale CAs. As in the large-scale case, we follow the two-step approach described in §2.1; in the first step, we parametrically estimate the distribution of competitors' bids using the model described in §4. Then, in the second step, we applied the two different markup models in the first-order conditions and compare the results. For the extended size-based markup model, bids with a winning probability above 10^{-3} were considered high winning probability (special) packages and were assigned a separate markup. On average, the extended size-based model uses only 35% of the markup variables of the full-dimension model.[19] The online appendix shows a scatter plot of the estimated per-meal markups from the two methods. Overall, the markup estimates of

[19] Because the relevance of each bid only depends on its winning probability and not on the markup specification, the set of relevant bids is the same for the two models, and therefore, they are comparable. Packages with winning probability below 10^{-4} were considered irrelevant. For each firm, the aggregate winning probability over these irrelevant bids is on average less than 1% of the firm's total winning probability. Hence, the effect of ignoring these irrelevant (with positive probability) bids was negligible.

the extended size-based markup model are similar to those obtained with the full-dimension model. The correlation between the markups is 0.982; their ratio is on average 1.003 with a standard deviation of 0.127. The estimates for the special packages are even closer to each other: the ratio is on average 0.998 with a standard deviation of 0.004 and the correlation is 0.999. This provides some support for the conjecture that grouping packages would have a negligible impact on the markups of the special packages, as discussed in §3.1.

We also note that this specification of the extended size-based markup model separates each size-group further into two subgroups with medium and low winning probability packages as described in the heuristic in §3.2. We observed that this additional refinement to the size-based markup model helped to improve the markup estimates (i.e., the estimates were closer to the estimates of the full-dimension model).[20]

To provide the performance metrics of a CA (winning bidders' profit margins and efficiency of the allocation), we need to estimate the total supplying costs of the winning firms in the CA and the efficient allocation. Notably, the estimated total supplying costs both in the winning CA and efficient allocations are very close between the two methods (differ by less than 0.1%).

We also compared the estimates of the full-dimension model with those of the *pure size-based* markup model, which does not isolate the high winning probability packages.[21] In this case, the ratio of the estimated per-meal markups from the two methods (with the full-dimension markup in the denominator) is on average 1.138 with standard deviation of 0.365. As expected, the pure size-based markup model results in significant bias in the estimated markups relative to the extended size-based model. In particular, not separating the high winning probability packages leads to overestimating the group markups. A partial explanation is that high winning probability packages tend to have larger estimated markups relative to the rest of the group.[22] Hence, following the discussion around Proposition 2, removing them from the group reduces the bias associated with grouping in an important way.

Finally, we also performed an experiment to examine the impact of package density on strategic markup reductions. As discussed in §3.1, economic theory predicts strategic markup reductions mainly driven by scale, and we did not incorporate explicitly the density effects in our extended size-based markup model (even though the separation of special packages may correct for it to some extent). To further justify this, we enriched the extended size-based markup model with a markup variable associated with a per-meal density measure of the package. The measure ranges between 0 and 1 and becomes larger as the package has more colocated units.[23] The estimates imply small markup reductions associated with density; they are on average 0.11% of the average bid price; this is an order of magnitude smaller compared to the markup reductions associated with the scale effect. Moreover, the extended size-based model with and without additional density parameters provide essentially the same markup estimates. The ratio of the two markup estimates is on average 0.999 with standard deviation of 0.063. This provides evidence that the density markup parameters do not play a significant role in the markup estimation.

Overall, this section provides evidence that our extended size-based model provides accurate approximations to the full-dimension model estimates, requiring significantly lower computational effort. In the small-scale CAs, the heuristic is an order of magnitude faster to run than the full-dimension model, producing similar estimates. In the large-scale CAs, the full-dimension model is computationally infeasible; we present the results using the extended size-based model heuristic in the next section.

6.2. Results for Large-scale CAs

The extended size-based markup model was used to estimate markups and costs for the package bids in the 2003 auction (later we report some results for 2005 as well; see also the online appendix).[24] In 2003, a total of 32 TUs in five regions were auctioned and 20 firms participated placing more than 2,000 bids per bidder on average.

[20] We used a winning probability of 6×10^{-4} to divide the groups into medium and low winning probabilities.

[21] Here, the pure size-based markup model also specifies markups per meal, including the volume of packages in the W matrix as described in §3.2.

[22] Let $\bar{m}_s^h$ and $\bar{m}_s^r$ be the *average* estimated per-meal markups of special packages with high winning probabilities (h) and the rest of the packages (r), respectively, for a particular firm and package size s in the full dimension model. Then, the ratio $\bar{m}_s^h/\bar{m}_s^r$ is on average 1.22 with a standard deviation of 0.37.

[23] The density measure is motivated by the density discount function used in §4. In particular, we tested two different density measures for robustness of the results and they both gave very similar results. We provide the details of these density measures and the results in the online appendix.

[24] After estimating the markup parameters, we numerically checked if the estimated markup variables locally optimize the expected profit. Note that to fully evaluate the local optimality of the markups, we need to estimate the Hessian matrices of the bidders' expected profit. However, estimating the Hessian matrices is computationally very intense, requiring an order of magnitude more computational time than estimating the markups. Instead, we checked the second-order derivatives of the bidders' expected profits with respect to each of the markup variables and they were all negative, consistent with the local optimality of the estimates.

Table 3 **Results from the Markup Estimation for Representative Firms of Different Winning Probability Levels for the 2003 Auction**

		Average markups for each package size						
Firm	Prob	1	2	3	4	5	6	Overall average
47	0.9193	22.64	15.07	12.14	7.98	7.54	7.19	9.88
36	0.6642	3.00	2.39	2.21	1.77	1.50	1.41	2.07
19	0.1578	0.81	0.82	0.84	0.79	0.72	0.71	0.79

Notes. "Prob" refers to the probability that the firm wins any package. The other columns show the average per-meal markups corresponding to each package size. The markups are shown as a percentage of the average bid price per meal (US$0.88).

After estimating the markups and costs of these firms, two groups of firms were identified based on each firm's *total winning probability*, that is, the firm's aggregate winning probability over all packages in the auction. The "competitive" group consists of 10 firms whose total winning probabilities are higher than 45%. Firms in the other group have very low winning probabilities (less than 2%) except for one with 16% of total winning probability.[25] In terms of markups, the competitive firms have markup estimates ranging from 1.2% to 18% of the average bid price with an average markup of 4.4% of the average bid price (US$0.88 per meal). The other firms have lower markups, resulting in an average markup over all firms of around 2.8% of the average bid price. Table 3 shows the average per-meal markup estimates for each package size (one through six units) for representative firms in three different levels of total winning probabilities. The estimates indicate that firms reduce their markups as the size of packages increases, showing that some portion of the discounts in package bids is due to markup reductions.

Firms submit hundreds to thousands of bids, and about 13% of them are relevant bids.[26] For the competitive firm group, the fraction of relevant bids is higher (22%). With the estimated markup and cost information of relevant bids, we are able to compute the total cost and markup of the CA allocation. The total procurement cost for the government was US$70.5 million per year and the total supplying cost for firms was US$67.2 million per year. This yields an average profit margin to winning firms of 4.8%. This level of profit margins is consistent with the Chilean government's estimate for this market. In addition, the Chilean government has their own estimates for the average TU costs, and these exhibit similar levels compared to our estimates. Both facts are reassuring.[27]

Finally, to compare results, we also performed the estimation for the 2005 auction, where 16 firms participated for 23 units. The results are consistent with the 2003 auction, both in the shape and level of the estimated markups. The total procurement cost amounts to US$ 53.4 million and the total supplying cost is US$51.5, which gives 3.5% of average profit margins to winning firms.

Next, we evaluate the cost synergies—cost savings from combining units together—implied by the estimates. Recall that our main objective is to determine what portion of the observed package discounts is due to cost synergies. Given the markup estimates, the per-meal cost of each package a submitted by firm f is given by $c_{af} = b_{af} - w_a\theta_f/v_a$, where θ_f is the markup vector estimated for that firm, b_{af} is the per-meal bid price placed by firm f for package a, and w_a is the ath row of package-characteristic matrix W used for bidder f. A direct calculation of the per-meal cost synergy in this package, denoted by s_a, can be computed from $s_a = \sum_{i\in a}(v_i/v_a)c_i - c_a$, where c_i is the point estimate for the cost of unit i. The cost synergies estimated directly from the cost estimates tend to increase as the size of packages grows. On average, they range from around 1.3% to 4.5% of the average bid price as the size of the package increases from two to eight, with an overall average synergy of 3.1%.

One disadvantage of estimating cost synergies in this direct way is that the synergies can only be computed for packages containing units whose single-unit bids are all relevant, which is not a representative sample of the bid population. To use a larger portion of the packages to estimate cost synergies, we run a regression similar to (7) but replacing the dependent variable b_{af} by c_{af}:

$$c_{af} = \sum_{i\in a} \xi_{if}\frac{v_i}{v_a} - g^{\text{scale}}(v_a, \gamma^{\text{scale}}_{k(f)}) - \sum_{c\in \text{Cl}(a)} g^{\text{density}}(v_c, \gamma^{\text{density}}_{k(f)})\cdot \frac{v_c}{v_a} + \varepsilon_{af}, \qquad (9)$$

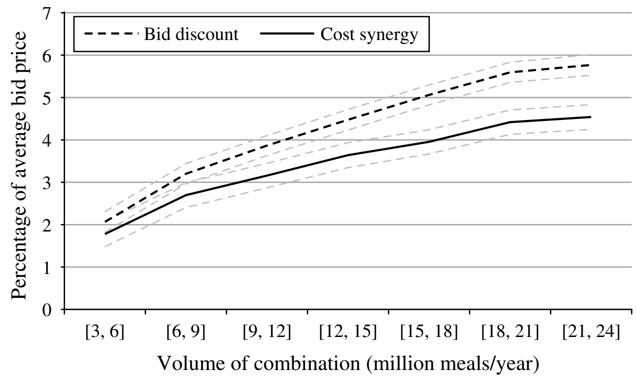
where again $k(f) \in \{L, O\}$ indicates one of the two firm types. This regression projects the estimated costs on scale and density synergies, and provides estimates of the costs for every unit, ξ_{if}, including those for which the single-unit bid was irrelevant. For relevant single-unit bids, ξ_{if} and c_{if} are quite close: their correlation is 0.993, the average absolute difference is

[25] In addition, from the 20 participating firms, there are two extreme firms with very competitive bids for which the estimated markups are unreasonably high and lead to negative costs for some packages. Despite their competitive prices, these firms did not win any units and were disqualified from the allocation process because of quality considerations. For these reasons, we omit them from our analysis hereafter.

[26] Packages with winning probabilities below 10^{-5} were considered irrelevant.

[27] Note that the government does not have estimates for TU costs for each firm, nor of the variance of these quantities. Moreover, they do not have reliable estimates for the level of cost synergies. Hence, the government's estimates are of course insufficient to evaluate the performance of the auction, which is the objective of this work.

Figure 1 Graph of Bid Discounts and Cost Synergies for the 2003 Auction



Notes. The dark dashed line shows the bid discounts estimated directly from the observed bid data using regression Equation (7) (the curve shows the total discounts, accounting for scale and density). The solid line shows the estimates of the total cost synergies (combining economies of scale and density) obtained from the structural estimation and regression Equation (9). Ninety-five percent confidence intervals are shown in gray dashed lines.

about 1%, and their ratio averages 1.00 with a standard deviation of 0.011. Hence, Equation (9) seems a reasonable approach to estimate cost synergies.

Figure 1 plots the total cost synergies and bid discounts as a function of package volume.[28] The results show that although there are some strategic markup reductions, most bid discounts (at least 75%) are actually explained by cost synergies. These synergies are quite significant and can be as large as 4.5% of the bid price on average.

The previous results suggest that in our application allowing package bidding may be appropriate: cost synergies are significant and account for most bid discounts vis-à-vis strategic markup reductions. Moreover, the overall markups that firms gain do not seem too large, resulting in a reasonable total procurement cost. Overall, our results suggest that the advantages of using package bidding (allow bidders to express cost synergies) may be larger than its disadvantages (the additional flexibility that firms can use to strategize and game the mechanism). In the remainder of this section, we use our estimates to provide sharper results concerning the performance of our CA. In particular, we study the allocative efficiency and procurement cost of the first-price sealed-bid CA, and compare it to alternative auction mechanisms.

6.3. Performance of the First-Price CA

In this section, we study the allocative efficiency of the 2003 CA.[29] The winning bidders' costs under the first-price CA allocation can be directly computed using the cost estimates obtained through the structural estimation. If we had the cost estimates for *all* possible packages, we could also calculate the efficient allocation, that is, the combination of package bids among all firms that achieve the minimum possible total cost. Unfortunately, our structural estimation method only identifies the costs of relevant bids, and the efficient allocation over this subset of combinations could overestimate the cost of the true efficient allocation that considers all possible packages.

To address this issue, we propose estimating the cost of irrelevant bid packages through an out-of-sample extrapolation based on Equation (9). However, the total number of feasible packages are in the order of millions and it is computationally infeasible to extrapolate to this entire set of feasible packages. Instead, we choose the set of packages on which at least one bidder placed a bid, which is in the order of 30,000 packages. We call this the *expanded package set*. Then, for each firm, we extrapolate costs to all packages in this expanded package set that are also in the set of feasible allocations. Although this is a small subset of all possible packages, it provides a reasonable approach to extend the set of bids observed in the data.

This out-of-sample extrapolation approach implicitly assumes that the selection of the bids in the irrelevant bid sample is independent of the costs of these units. Recall that irrelevant bids include bids that were not submitted by the bidder. Hence, in our application, it could be possible that the sample selection of irrelevant bids is related to costs; for example, bidders are likely to bid on the subset of combinations where they are more competitive, so that higher-cost combinations are not submitted. If this is the case, then our cost extrapolation procedure could lead to a cost estimate of the efficient allocation that is lower than the true one, so that we could *overestimate the true efficiency loss* of the first-price CA.

Recall that in 2003, the bidders' supplying costs given by the auction allocation were equal to US$67.2 million per year. The efficient allocation that minimizes the total supplying cost among the feasible allocations over the set of relevant bids is equal to US$66.7 million per year, implying an efficiency loss of 0.65%. When considering the expanded package set, the total supplying cost of the efficient allocation goes down to US$66.2 million per year, with an efficiency loss of the first-price CA of 1.5%.[30] The

[28] There are two small firms whose estimated cost synergies are significantly different from the rest of the firms, and they are not included in the figure.

[29] The results for the 2005 auction are similar and consistent with those for the 2003 auction. We provide the counterfactual results for the 2005 auction in the online appendix.

[30] It is worth noting that the first-price CA tends to identify the most cost-efficient firms in the different geographical regions. More specifically, there are nine firms in the CA allocation and 10 firms in the efficient allocation; the majority of them—seven firms—appear in both allocations. Two firms are allocated the exact same set of packages in both cases and other firms win packages that contain many overlapping units or units from the same geographical regions.

efficiency loss is arguably low. We believe this result is essentially driven by the high level of competition in the auction; there is a reasonable number of firms and most firms compete in all units and submit many package bids. For this reason, firms do not seem to have enough market power to significantly harm efficiency by using the flexibility that package bidding allows with strategic motivations.

Although the efficiency loss is overall evidently low, a few firms indeed engage in strategic markup reductions that are consistent with the economic arguments provided in §1. For example, there is one firm winning eight units in the CA allocation, that essentially leveraged its cost advantage in some units to win another unit for which it was not the cost-efficient firm. If this firm was forced to just win its cost-efficient bundle, the loss of 0.65% over relevant bids is significantly reduced. In summary, the high efficiency and relatively small profit margins for firms (around 5% as presented in §6.2) achieved by the school meals CA suggest that it is a reasonable mechanism for the procurement of this public service.

6.4. Evaluation of the Vickrey-Clarke-Groves Mechanism

Although our previous results support that using the first-price CA in our application seems appropriate, it is also useful to compare the performance with alternative auction mechanisms. In this section, we perform a counterfactual to compare the performance with the Vickrey-Clarke-Groves (VCG) mechanism, which generalizes the second-price auction in CA settings. It is well known that for the VCG mechanism truthful bidding is a dominant strategy (i.e., bidders report their true costs), and hence VCG achieves the fully efficient allocation.[31] For this reason, in a public procurement setting like our application, VCG could potentially be an attractive alternative.[32] However, despite these advantages, the VCG mechanism is rarely applied in practice and has been criticized for other numerous drawbacks. In particular, Ausubel and Milgrom (2006) have shown that in the face of complementarities, the VCG procurement costs can be prohibitively high. Hence, it is on itself interesting to see how VCG performs in real-world applications.

In our analysis, we use the same set of extrapolated bids as in §6.3 as the bids (costs) that bidders would report in the VCG mechanism. We know VCG achieves the efficient allocation, which was previously computed in §6.3. From the bids, we can compute the individual VCG payments to the winning bidders, and by summing them, we obtain the VCG procurement cost. As seen in the previous section, the total annual procurement cost in the 2003 first-price CA is US$70.5 million. The total annual procurement cost under the VCG mechanism is US$70.3 million, which is about 0.32% cheaper than the first-price CA.

The result is at odds with the theoretical literature mentioned above describing the pitfalls of VCG; in our application, VCG achieves payments comparable to the first-price CA and induces a reasonable procurement cost. We believe this result is driven by the significant amount of competition introduced by the large number of package bids submitted by firms. In this case, a winning bidder is not that relevant; if its bids are eliminated, there is another allocation that achieves costs close to the minimum-cost allocation, leading to reasonably low VCG payments.[33] In contrast, in the examples provided by Ausubel and Milgrom (2006), competition is limited, resulting in high VCG payments. Hence, VCG should achieve reasonable procurement costs in settings with a reasonable amount of bidders that are able to submit many package bids with significant coverage of all the units in the auction. The latter should be expected when every unit is attractive to many firms and it is relatively effortless for a bidder to evaluate its costs for different packages.

7. Conclusions

In this paper, we develop a structural estimation approach that allows evaluating the performance of large-scale first-price CAs. An important methodological contribution of our work is to introduce a restricted markup model in which bidders are assumed to determine their markups based on a reduced set of package characteristics. The main advantage of this approach is that it reduces the computational burden of the structural approach so that it can be applied to large-scale CAs.

We effectively apply our structural estimation approach to the large-scale Chilean school meals CA. We find that cost synergies in this auction are significant and the current CA mechanism, which allows firms to express these synergies through package bidding, seems appropriate. In particular, the current CA achieves high allocative efficiency and a reasonable procurement cost. We believe this is the first empirical analysis documenting that a CA performs well in a real-world application.

[31] In the online appendix, we provide details about VCG and its payment rule.

[32] It could also be useful to compare the performance with other alternative auction mechanisms (for which bidders are not truthful). However, such counterfactuals are limited as they require computing the bidding strategies played in equilibrium by the bidders. Unfortunately, equilibrium results for most of the multiunit auction mechanisms that are used in practice are at best rare.

[33] In the online appendix, we also discuss how this result relates to the *closeness* of the VCG payoffs to the core of the transferable utility cooperative game played among the bidders and the auctioneer.

More broadly, our results highlight the importance of the joint consideration of the firms' operational cost structure and their strategic behavior for the successful design of a CA. Moreover, despite the practical use of CAs, its econometric analysis has been limited because of its complexity. Even though our method may need modification to accommodate different pricing and auction rules, we believe it can be a useful starting point to reduce the complexity of econometric analysis in other large-scale settings. In this way, we hope that this research agenda enhances the understanding of the performance of CAs and thereby provides insights to improve their design.

Acknowledgments

The authors thank Rafael Epstein for generously sharing the data and the valuable insights he provided. The authors are also thankful for the detailed comments provided by Greg Lewis and Cristián Martínez. Serguei Netessine (the department editor), an anonymous associate editor, and anonymous referees provided constructive feedback that significantly improved this paper. Special thanks to Alfonso Acuña, who provided valuable insights on the bidding process from a practitioner's perspective. The authors had helpful conversations with Heski Bar-Isaac, Lanier Benkard, Estelle Cantillon, Ignacio Esponda, Awi Federgruen, Jeremy Fox, Kate Ho, Ali Hortaçsu, Jakub Kastl, Kei Kawai, Saravanan Kesavan, Paul Milgrom, Ariel Pakes, Martin Pesendorfer, Margaret Pierson, Mar Reguant, Michael Riordan, Richard Steinberg, and Assaf Zeevi, as well as seminar participants at the National Bureau of Economic Research Industrial Organization Summer Institute, Workshop on Empirical Research in Operations Management, Manufacturing and Service Operations Management Special Interest Group Conference, and various other conferences and institutions. The authors thank Daniel Yung for exemplary research assistance, as well as the Social Enterprise Program, the Center for International Business Education and Research, and the W. Edwards Deming Center at Columbia Business School for financial support. Marcelo Olivares also thanks Instituto Sistemas Complejos de Ingenieria for financial assistance.

References

Akşin Z, Ata B, Emadi SM, Su C-L (2013) Structural estimation of callers' delay sensitivity in call centers. *Management Sci.* 59(12): 2727–2746.

Allon G, Federgruen A, Pierson M (2011) How much is a reduction of your customers' wait worth? An empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing Service Oper. Management* 13(4):489–507.

Athey S, Haile PA (2006) Empirical models of auctions. Blundell R, Newey WK, Persson T, eds. *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, Vol. II (Cambridge University Press, New York), 1–45.

Ausubel LM, Milgrom P (2006) The lovely but lonely Vickrey auction. Cramton P, Shoham Y, Steinberg R, eds. *Combinatorial Auctions* (MIT Press, Cambridge, MA), 17–40.

Baranov O (2010) Exposure vs. free-riding in auctions with incomplete information. Working paper, University of Colorado at Boulder, Boulder.

Bichler M, Davenport A, Hohner G, Kalagnanam J (2006) Industrial procurement auctions. Cramton P, Shoham Y, Steinberg R, eds. *Combinatorial Auctions* (MIT Press, Cambridge, MA), 593–612.

Cantillon E, Pesendorfer M (2006) Combination bidding in multi-unit auctions. Working paper, Université Libre de Bruxelles, Brussels, Belgium.

Caplice C, Sheffi Y (2006) Combinatorial auctions for truckload transportation. Cramton P, Shoham Y, Steinberg R, eds. *Combinatorial Auctions* (MIT Press, Cambridge, MA), 539–572.

Chu CS, Leslie P, Sorensen A (2011) Bundle-size pricing as an approximation to mixed bundling. *Amer. Econom. Rev.* 101(1): 263–303.

Cramton P, Shoham Y, Steinberg R (2006) *Combinatorial Auctions* (MIT Press, Cambridge, MA).

Elmaghraby W, Keskinocak P (2004) Combinatorial auctions in procurement. Harrison T, Lee H, Neale J, eds. *The Practice of Supply Chain Management: Where Theory and Application Converge* (Springer, New York), 245–258.

Epstein R, Henríquez L, Catalán J, Weintraub GY, Martínez C (2002) A combinational auction improves school meals in Chile. *Interfaces* 32(6):1–14.

Fox JT, Bajari P (2013) Measuring the efficiency of an FCC spectrum auction. *Amer. Econom. J.: Microeconomics* 5(1):100–146.

Glynn P (1989) Optimization of stochastic systems via simulation. MacNair EA, Musselman KJ, Heidelberger P, eds. *Proc. 1989 Winter Simulation Conf.*, 90–105.

Guerre I, Perrigne I, Vuong Q (2000) Optimal nonparametric estimation of first-price auctions. *Econometrica* 68(3):525–574.

Hendricks K, Porter RH (2007) An empirical perspective on auctions. Armstrong M, Porter R, eds. *Handbook of Industrial Organization*, Vol. 3 (North-Holland, Amsterdam), 2073–2143.

Hohner G, Rich J, Ng E, Reid G, Davenport A, Kalagnanam J, Lee H, An C (2003) Combinatorial and quantity-discount procurement auctions benefit Mars, Incorporated and its suppliers. *Interfaces* 33(1):23–35.

Hortaçsu A, McAdams D (2010) Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the Turkish Treasury auction market. *J. Political Econom.* 188(5): 833–865.

Kastl J (2011) Discrete bids and empirical inference in divisible good auctions. *Rev. Econom. Stud.* 78(3):974–1014.

Krasnokutskaya E (2011) Identification and estimation of auction models with unobserved heterogeneity. *Rev. Econom. Stud.* 78(1): 293–327.

Li J, Granados N, Netessine S (2014) Are consumers strategic? Structural estimation from the air-travel industry. *Management Sci.*, ePub ahead of print March 20, http://dx.doi.org/10.1287/mnsc.2013.1860.

McDuff DF (2003) Analyzing package bidding in the FCC auction no. 31: Upper 700 MHz band. Accessed December 17, 2013, http://cramton.umd.edu/econ415/project-fcc-package-auctions.pdf.

Milgrom P (2000) Putting auction theory to work: The simultaneous ascending auction. *J. Political Econom.* 108(2):245–272.

Olivares M, Terwiesch C, Cassorla L (2008) Structural estimation of the newsvendor model: An application to reserving operating room time. *Management Sci.* 54(1):41–55.

Olivares M, Weintraub GY, Epstein R, Yung D (2012) Combinatorial auctions for procurement: An empirical study of the Chilean school meals auction. *Management Sci.* 58(8):1458–1481.

Paarsch HJ, Hong H (2006) *An Introduction to the Structural Econometrics of Auction Data* (MIT Press, Cambridge, MA).

Reguant M (2011) The welfare effects of complementary bidding mechanisms: An application to electricity markets. Working paper, Stanford University, Stanford, CA.

Sandholm T, Levine D, Concordia M, Martyn P, Hughes R, Jacobs J, Begg D (2006) Changing the game in strategic sourcing at Procter & Gamble: Expressive competition enabled by optimization. *Interfaces* 36(1):55–68.

Wooldridge JM (2002) *Econometric Analysis of Cross Section and Panel Data* (MIT press, Cambridge, MA).

CORRECTION

In this article, "Measuring the Performance of Large-Scale Combinatorial Auctions: A Structural Estimation Approach," by Sang Won Kim, Marcelo Olivares, and Gabriel Y. Weintraub (*Management Science*, 2014, Vol. 60, No. 5, pp. 1180–1201), the following sentence, which appears on page 1181, was corrected to read as follows: "If the bidder wins the package, it will lead to an inefficient allocation in which a unit is not served by the lowest-cost supplier."