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

Sang Won Kim ${ }^{\dagger}$, Marcelo Olivares ${ }^{\ddagger}$, and Gabriel Y. Weintraub ${ }^{\S}$

May 2, 2012


#### Abstract

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, reducing the performance of the auction. In this paper, we develop a structural estimation approach for large-scale first-price CAs. We use bidding data to estimate the firms' cost structure and evaluate the performance of the auction in terms of the cost efficiency of the allocation and payments to the bidders. 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 where markups of each package bid are chosen based on a reduce set of package characteristics. We apply our method to the Chilean school meals auction, in which the government procures half a billion dollar 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, and the current CA mechanism achieves high allocative efficiency and a reasonable procurement cost. We also perform counterfactuals to compare the performance of the current CA with alternative mechanisms such as VCG.


Keywords: combinatorial auctions, procurement, empirical, structural estimation, auction design, public sector applications.

[^0]
## 1 Introduction

In many important procurement settings suppliers face cost synergies driven by economies of scale or density. 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 this type 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 theory. In fact, these multi-unit auctions have been successfully implemented in many applications, including the procurement of school meal services, bus routes, electricity, transportation services, and inputs in private firms, as well as in non-procurement settings, such as the auctions for wireless spectrum run by the Federal Communications Commission (FCC).

In this paper we introduce a structural estimation approach to empirically analyze the performance of first-price combinatorial auctions (CAs), a multi-unit mechanism that allows bidders to submit separate bids for different combinations or packages of units (see Cramton et al. (2006) for an overview on CAs). We address a central design question: what is the impact that allowing package bidding via a CA has on performance. Our analysis considers two performance measures: (1) efficiency, which relates the actual bidders' costs incurred in the CA allocation relative to the minimum possible cost allocation any subset of suppliers could achieve; and (2) optimality, which relates to the total expected payments to bidders by the auctioneer.

There are two countering effects that can affect the performance of a CA. The main advantage of package bidding is that it allows bidders to express cost synergies in their bids. In contrast, 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, makes the bidders less aggressive in expressing the cost savings of supplying multiple units. Allowing package bidding eliminates this risk, potentially leading to more efficient outcomes and lower procurement costs.

On the other hand, allowing package bids can also hurt the performance of a first-price CA. As pointed out by Cantillon and Pesendorfer (2006) and Olivares et al. (2011), 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 $A$ (for which the bidder is the cost efficient provider) into another unit $B$ (for which the bidder is not the efficient provider), by submitting a discounted package bid for $A$ and $B$ to try to win both units. If the bidder wins the package it will lead to an inefficient allocation where a unit ends up being served by a supplier with a higher cost. In addition, package bidding can also lead to a free-riding problem (also known as 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 and thereby higher payments for the auctioneer (Milgrom, 2000).

Due to the aforementioned trade-off between cost synergies and strategic behavior introduced by package bidding, it is important to analyze the actual performance that CAs have in practice. If cost synergies are strong and the incentives for the types of strategic behavior mentioned above are weak, then a CA should achieve a good performance relative to auction mechanisms that preclude package bidding. On the other
hand, if the strategic motivations alluded above are strong and cost synergies are weak, package bidding may hurt the efficiency and optimality of the auction. Hence, understanding the performance of a CA requires knowledge of cost synergies, as well as the incentives that drive strategic behavior. Unfortunately, existing theory is not conclusive on how large these incentives are in a specific application. Moreover, cost information is typically private and sensitive information of the bidders. Thus motivated, the objective of this paper is to provide an empirical methodology that can be used to evaluate the performance of a first-price sealed-bid CA using bid data. We apply our method to the Chilean school meals combinatorial auction that we describe below.

To measure the performance of a CA, it is essential to identify bidders' supplying costs, which are not directly observable in the bid data. In the context of the application we study, previous work by Olivares et al. (2011) provides evidence of significant package discounts (see Figure 1). Note that even if the bid prices of packages decrease as the package sizes increase, these discounts cannot be directly interpreted as cost synergies. They could also be driven by markup adjustments motivated by the types of strategic behavior alluded above. Although Olivares et al. (2011) provide further suggestive evidence of cost synergies in this application, their approach does not provide direct estimates of the suppliers' costs, which are needed to evaluate the performance of the auction. In addition, their approach does not allow to perform counterfactuals. Other reduced form approaches suffer from the same limitations (see Ausubel et al. (1997), Gandal (1997), and Moreton and Spiller (1998)).

As an alternative to this reduced form approach, we propose a structural estimation approach which directly identifies the bidders' costs using actual bid data. In particular, our structural method disentangles whether the discounts observed in bids are driven by cost synergies or strategic markup adjustments. Our method is based on the seminal work of Guerre et al. (2000) for single-unit auctions that was later extended by Cantillon and Pesendorfer (2006) to a CA setting with a small number of units. The main idea behind this structural approach is to use the first-order conditions from the bidder's profit maximization problem to find the imputed costs that would rationalize the bids observed in the data. Because the bidder's problem involves beliefs about the competitors' bidding behavior, this approach also requires estimating a statistical distribution of the competitors' bids.

In the large-scale CA that we analyze in this paper - where each bidder submits in the order of hundreds or thousands of bids - a direct application of the Cantillon and Pesendorfer (2006) method is not possible due to the large number of decision variables in the bidder's profit maximization problem. We develop a novel approach to overcome this issue, assuming a "simplified" version of the bidder's problem where the markups charged on the package bids are chosen based on a reduced set of package characteristics. With this simplification, the bidder's problem becomes computationally and econometrically tractable so that the structural approach can be effectively applied to large-scale CAs. Recall, however, that the main objective of the structural approach is to identify the cost structure - which is a primitive in the structural model separately from the markups, which is chosen strategically by the bidders. Therefore, in using our proposed approach it is important to allow for sufficient flexibility in the markup specification so that strategic markup adjustments are not overly restricted in this simplified bidder's problem. We provide detailed guidelines on how to do this in the estimation.

We expect that our approach, based on pricing package characteristics, can be used in several real-
world large-scale auctions. In particular, we effectively apply our method to the Chilean school meals CA (see Epstein et al. (2002) for a detailed description of the auction). This application fits well within the class of large-scale CAs: each auction has about 30 units and firms submit hundreds or even thousands of bids. This CA has a single-round sealed-bid first-price format. The auction is used by the Chilean government to allocate contracts among private catering firms to provide breakfast and lunch for 2.5 million children daily in primary and secondary public schools during the school year. In a developing country where about 14 percent 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. The CA, one of the largest state-run auctions in Chile, was used for the first time in 1999 and has been used every year since its inception awarding more than $\$ 3$ billion of contracts. Although this application has been praised for bringing transparency and lowering the procurement costs of a high social impact public service, a detailed performance analysis of the CA format has not been conducted.

Our results show that for the Chilean auction, cost synergies are significant, amounting up to $6 \%$ of the cost. Roughly $75 \%$ of the discounts observed in the bid data arise from cost synergies (the rest is due to strategic markup adjustments). In part due to this large cost synergies, the CA achieves a strikingly high efficiency, with an actual cost allocation only $1 \%$ higher than the minimum-cost allocation. The results also show that while economies of scale (mostly generated by volume discounts in input purchases) are larger than economies of density (arising from common logistics infrastructure used to supply nearby units), they are both important in the firms' operational cost synergies. Finally, the estimated markups are on average around $5 \%$, suggesting that the CA induces a reasonable amount of competition among the suppliers. The level of markups coincides with anecdotal evidence provided by the Chilean government. Going back to our initial motivating auction design question, overall, our results suggest that package bidding and running a CA seems appropriate in our setting. Further, the results suggest that the bidding language should allow bidders to express both economies of scale and density.

Once we estimate the cost structure we can also perform other useful counterfactuals. One important consideration the government has when running these auctions, which arises frequently in other settings with synergies, is how to promote diversification and competition among bidders. In the Chilean auction, the government imposes market share restrictions for bidders in the CA to promote long-run competition. The cost estimates provided by the structural estimation can be used to evaluate the efficiency loss due to these constraints. We find that the efficiency loss is very small, around $1 \%$. The main reason for this result is that cost synergies get practically exhausted at the point where the market share constraints become binding.

An important practical motivation to use a first-price rule in applications of CAs is that a Vickrey-Clarke-Groves (VCG) mechanism, that is known to be truthful and efficient, can lead to excessively large payments and other undesirable properties in the presence of synergies (Ausubel and Milgrom, 2006). We conduct a counterfactual experiment to compare the total payments of the first-price sealed-bid CA against a VCG mechanism. Finding the counterfactual total payment of VCG requires identifying the efficient allocation, which can be computed from the cost estimates obtained via the structural estimation. Interestingly, and contrary to the theoretical results mentioned above, we find that in our application the total VCG payment is quite reasonable and very close to the first-price CA payment. We believe this result is driven by the significant amount of competition introduced by the large number of package bids submitted by firms.

Our work is related to other structural estimation papers in auctions (see Athey and Haile (2006), Hen-
dricks and Porter (2007), and Paarsch and Hong (2006) for good surveys). Most notably, Reguant (2011) uses a first-order-conditions approach to structurally estimate a model of the day-ahead wholesale electricity market in Spain. After estimating the cost structure of bidding firms, she performs counterfactuals to determine the welfare effects of "complex bids"; a specific bidding mechanism that allow companies to express cost complementarities of operating across different hours of a day. In addition, Fox and Bajari (2011) use an estimator based on a pairwise stability condition to estimate complementarities in an FCC spectrum auction, which is run in an ascending auction format without package bidding.

There has also been an important literature studying multi-unit auctions of homogeneous goods. We describe a sample of these papers here. Hortaçsu and McAdams (2010) develop a structural estimation approach for the Turkish treasury auction and uses the estimates to compare the performance of a uniform price and a Vickrey auction. Kastl (2011) studies how to use a structural approach when bidders submit discrete bid points rather than continuous downward sloping demand functions with data from the Czech Treasury auction. Finally Chapman et al. (2005) develop a framework to measure best-response violations in multi-unit, sealed-bid, discriminatory-price auctions run by the Bank of Canada to manage excess cash reserves.

Our work is also related to the growing literature in operations management that uses structural estimation. Olivares et al. (2008) develop a structural approach to impute the cost of overage and underage of a newsvendor, which is applied to the reservation of operating room time by an hospital. Allon et al. (2011) conduct a structural estimation to measure the implicit waiting cost of customers in the fast food industry. Similarly, Aksin-Karaesmen et al. (2011) estimate customer waiting costs but develop a dynamic structural model to explain customer abandonments in a bank's call center. Li et al. (2011) also model consumer's forward looking behavior through a dynamic structural model, using data from the airline industry. 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 structural estimation methods have not been used.

The rest of the paper is structured as follows. Section 2 develops a structural estimation framework to estimate the primitives of first-price sealed-bid CAs and proposes our structural model for large-scale CAs. Section 3 provides a description of the Chilean auction for school meals and our data set. Section 4 describes the details of our estimation method for the Chilean school meals auction and reports the estimation results. We evaluate the current auction format through efficiency analysis and perform counterfactual analysis in Section 5. Section 6 describes our main conclusions.

## 2 Structural Estimation Framework

This section develops a structural estimation framework to estimate the primitives of large-scale first-price single-round sealed-bid CAs. First, in Section 2.1 we describe a general structural estimation framework for CAs. A similar approach has been successfully implemented by Cantillon and Pesendorfer (2006) (hereon CP ) to estimate small London bus route CAs. This approach is itself inspired by the pioneering work by Guerre et al. (2000) for single-unit auctions. However, there are limitations of using this approach when the number of units and possible packages grows, which is common for many CAs in practice. Section 2.2 describes our contribution to address these limitations, which basically relies on assuming a simplified bidding
strategy that reduces the complexity of the bidder's problem, making the structural estimation feasible for large-scale CAs. Section 2.3 provides guidelines on how to implement our structural estimation approach and Section 2.4 discusses important identification issues.

### 2.1 A Structural Estimation Approach to First-Price Sealed-Bid Combinatorial Auctions

We begin by describing a structural estimation approach to first-price single-round sealed-bid CAs. The approach is similar to the one introduced by CP .

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 non-empty 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 possible packages. Let $b_{a f}$ denote the bid price asked by bidder $f$ to supply package $a$, and $b_{f}=\left\{b_{a f}\right\}_{a \in \mathcal{A}}$ 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:

$$
\begin{array}{ll}
\min & \sum  \tag{1}\\
\text { s.t. } & b_{a f} x_{a f} \\
& x \in X, \quad x_{a f}=\{0,1\}, \forall a \in \mathcal{A}, f \in F,
\end{array}
$$

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

The winner determination problem minimizes the total procurement costs 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 Section 3.2 and the online appendix we provide more details on the mathematical integer program that solves the winner determination problem in the context of our specific empirical application.

The structural estimation approach requires assumptions on the bidders' information structure and bidding behavior in order to identify costs. We make the following assumption.

Assumption 2 (Bidders' Costs). Bidders have independent private costs. In particular, given an auction, each bidder gets an independent random draw of a cost vector $c_{f}=\left\{c_{a f}\right\}_{a \in \mathcal{A}}$, where $c_{a f}$ 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. We make the following assumption on bidders' strategies.

Assumption 3 (Strategies). Bidders are risk-neutral and play pure bidding strategies. A bidder's strategy is a function $b_{f}: \Re_{+}^{A} \mapsto \Re_{+}^{A}$ 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 asymmetric 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 $G_{f}\left(b_{f}\right)=\left\{G_{a f}\left(b_{f}\right)\right\}_{a \in \mathcal{A}}$, where $G_{a f}\left(b_{f}\right)$ is the probability 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:

$$
\begin{equation*}
\max _{b \in \Re^{A}}(b-c)^{T} G(b), \tag{2}
\end{equation*}
$$

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. Whenever the context is clear, we omit the subscript $f$ to simplify the notation.

To formulate the optimization problem above, a bidder needs to form expectations about the bidding behavior 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 $f^{\prime}$ uses a bidding strategy $b_{f^{\prime}}(\cdot)$, the bids of bidder $f^{\prime}$ are random from bidder $f^{\prime}$ 's perspective; they correspond to the composition $b_{f^{\prime}}\left(c_{f^{\prime}}\right)$, where $c_{f^{\prime}}$ is the random cost vector for bidder $f^{\prime}$. Assumption 4, described next, formalizes this. Assumptions 1, 2, 3, and 4 are kept throughout the paper.

Assumption 4 (Bid Distributions). a) Consider a given auction and any bidder f. From its perspective, assume that competitors' bid vectors $b_{f^{\prime}}$ are drawn independently from distributions $H\left(\cdot \mid Z_{f^{\prime}}\right)$, where $Z_{f^{\prime}}$ is a vector of observable characteristics of bidder $f^{\prime}$. These distributions are common knowledge among bidders and induce the correct vector of winning probabilities $G_{f}\left(b_{f}\right)$, for all $b_{f}$, given the competitors' strategies and the cost vectors' probability distributions.
b) In addition, assume that for all bidders $f \in F, H\left(\cdot \mid Z_{f}\right)$ has a continuous density everywhere.

Note that the independence part of the assumption is consistent with Assumptions 2 and 3. Also note that Assumption 4 captures all the relevant uncertainty faced by the bidder when solving (2). In particular, for a given bid vector $b_{f}$ submitted by bidder $f$, the competitors' bid distributions $\left\{H\left(b_{f^{\prime}} \mid Z_{f^{\prime}}\right)\right\}_{f^{\prime} \neq f}$ and the allocation rules given by the winner determination problem uniquely determine winning probabilities of each bid $G_{a f}\left(b_{f}\right)$.

Previous work, like CP and Guerre et al. (2000), assume 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-item auction studied in Guerre et al. (2000) this is well justified; under mild conditions a unique symmetric BNE always exist. However, there are no theoretical results that guarantee existence, uniqueness, nor characterization of equilibrium for a CA. Hence, we make slightly weaker assumptions, but that still allow us to develop a structural estimation approach.

More specifically, note that assuming BNE play imposes two conditions: (i) bidders correctly anticipate the strategy of their competitors, and therefore correctly estimate the vector of winning probabilities given their own bids; and (ii) for each bidder, given its costs and the winning probabilities function, the bidder selects a bid vector that maximizes its expected profit. While Assumption 4.(a) is weaker than condition (i), it imposes the same restriction over bidders' beliefs that we use in our structural estimation approach: bidders in the auction can correctly anticipate their winning probabilities. We also make a weaker assumption relative to the aforementioned condition (ii) imposed by BNE: we will only assume that bidders select a bid vector that satisfies the necessary first-order conditions of the expected profit maximization problem, and these are not sufficient for optimality in a CA. We come back to this point in the sequel.

Assumption 4.(b) guarantees the differentiability of the winning probability vector $G(\cdot)$ that is needed to use the first-order conditions for estimation, as we formalize in the next lemma. Note that this assumption is over the bids' distribution, that is endogenous 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 equilibrium in CAs does not allow us to follow this approach.

Proposition 1. Consider a given auction. For every bidder, the winning probability vector $G(b)$ is continuous and differentiable, for all b.

The proof of this proposition as well as all other proofs are provided in Online Appendix B. For a given bidder, the necessary first-order conditions of the optimization problem (2) are given by the following vector equation:

$$
\begin{equation*}
c=b+\left\{\left[\mathcal{D}_{b} G(b)\right]^{T}\right\}^{-1} G(b), \tag{3}
\end{equation*}
$$

where $\mathcal{D}_{b}$ refers to the Jacobian matrix operator with respect to the variable vector $b$ so that the $i j^{\text {th }}$ element is $\left[\mathcal{D}_{b} G(b)\right]_{i j}=\frac{\partial}{\partial b_{j}} G_{i}(b)$. The Jacobian is a square matrix which can have non-zero off-diagonal elements because packages of the same bidder compete against each other. Note that for a given auction there is one first-order condition vector equation per bidder; these equations are the basis to identify the cost vectors of each bidder, as we now explain.

Notice that, for a given bidder, the right-hand side of equation (3) only depends on the observed bid vector $b$, the winning probabilities $G(b)$ and its derivatives. By Assumption 4, the vector of winning probabilities $G(b)$ must be consistent with the actual auction play observed in the data, and therefore can be potentially estimated using bidding data from all bidders. For example, Guerre et al. (2000) estimate the distribution $G(\cdot)$ (and its derivative), which in a single-unit procurement auction corresponds to the tail distribution of the minimum bid, using a non-parametric approach.

In a CA setting, $G(\cdot)$ is a vector of probabilities, which complicates its estimation. A possible approach to estimate $G(\cdot)$ is to parametrically estimate the bid distribution of competing bidders $\left(H\left(\cdot \mid Z_{f}\right)\right.$ in Assumption 4) using bidding data. Note that this is a highly dimensional distribution so the parametric as-
sumptions will be important to make the estimation tractable. In Section 4.1 we provide more details about a parsimonious, yet flexible parametric description of this distribution in the context of our application.

Note also that previous structural approaches usually use a cross section of auctions for the estimation of $G(\cdot)$, assuming that in all auctions the same equilibrium is being played. In our approach we conduct an auction-by-auction estimation which does not require the latter assumption. Also, our estimates are less susceptible to unobserved heterogeneity across auctions, that is, to the effect of characteristics observed by the bidders but not by the econometrician that we do not control for in $\left\{Z_{f}\right\}_{f \in F}$. As we discuss later, for a given auction, we exploit the large number of units and packages in our application to parametrically estimate the distribution of competitors' bids.

Using the distribution of competitors' bids, one can use simulation to estimate the winning probabilities by sampling competitors bids from these distributions and solving the winner determination problem repeatedly. Derivatives could then be computed using a finite difference method. In fact, CP uses this method and replaces the estimates of $G(b)$ and $\mathcal{D}_{b} G(b)$ together with the observed bid vector $b$ in equation (3) to obtain an estimate of $c$ for each firm. They were able to effectively use this approach in auctions of at most 3 units.

Even if one was able to parametrically estimate the distribution of competitors' bids, there is an important limitation of the previous approach in larger-scale CAs: the dimensionality of the optimization problem (2) increases exponentially with the number of units. For example, in our application there are millions of possible packages and bidders submit in the order of hundreds or thousands of bids. Let us revise the first-order condition (3) in this context. First, for a given bidder, we need to estimate hundreds or thousands of winning probabilities. As the number of bids submitted by a bidder increases, the winning probability of each bid is likely to become very small and the simulation error of these low-probability events becomes large. Moreover, equation (3) requires taking derivatives over a large number of variables; simulation error for these quantities may be even larger. Hence, computation of $G(b)$ and $\mathcal{D}_{b} G(b)$ 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. One approach to simplify the bidders' problem to make it more amenable for analysis is to reduce the set of decision variables. The next section describes a structural model which incorporates this simplification.

### 2.2 The Characteristic-Based Markup Approach for CAs

Our model is based on the bidder's problem (2), which we refer to as the full-dimension problem in the sense that the bidder chooses every bid price. As mentioned above, the main complication of using this model in a large-scale CA is that the dimension is too large. In what follows, we present an approach to reduce the dimensionality of the problem. In particular, we develop a structural estimation approach that imposes 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 the first-order condition (3) can be re-written as $b=c+\left(-\left\{\left[\mathcal{D}_{b} G(b)\right]^{T}\right\}^{-1} G(b)\right)$, so that the bid is a cost plus a markup. Hence, we can view the full-dimension problem as choosing a markup for each package. We propose instead that this markup is specified by a reduced set of package characteristics.

Specifically, let $w_{a}$ be a row vector of characteristics describing package $a$, with dimension $\operatorname{dim}\left(w_{a}\right)=$ $d$ much smaller than $A$. The markup for package $a$ is given by the linear function $w_{a} \theta$, where $\theta$ is a (column) vector of dimension $d$ specifying the markup associated with each package characteristic. Instead of choosing the markup for each package, the bidder now chooses $\theta$ - the set of markups associated with each of these reduced set of package characteristics. Let $W \in \Re^{A \times d}$ be a matrix containing the characteristics of all packages, so that the $a^{\text {th }}$ 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 $\theta$ is a d-dimensional decision vector chosen by the bidder.

Note that different bidders can adopt different $W$ matrices. Under this assumption, the bidder's optimization problem becomes:

$$
\begin{equation*}
\max _{\theta \in \Re^{d}}(W \theta)^{T} G(W \theta+c), \tag{4}
\end{equation*}
$$

whose first-order conditions yield:

$$
\begin{equation*}
\left[\mathcal{D}_{\theta} W^{T} G(W \theta+c)\right]^{T} \theta=-W^{T} G(W \theta+c) . \tag{5}
\end{equation*}
$$

Here again the $i j^{t h}$ element of the Jacobian matrix above is $\left[\mathcal{D}_{\theta} W^{T} G(W \theta+c)\right]_{i j}=\frac{\partial}{\partial \theta_{j}}\left[W^{T} G(W \theta+c)\right]_{i}=$ $\frac{\partial}{\partial \theta_{j}} W_{i}^{T} G(W \theta+c)$, where $W_{i}$ is the $i^{\text {th }}$ column of matrix $W$. Re-arranging and replacing terms, we can solve for the decision vector $\theta$ as follows:

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

As in Guerre et al. (2000) and Cantillon and Pesendorfer (2006), 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 vector equation per bidder and, for each bidder, under Assumption 5, the cost is given by $c=b-W \theta$. Hence, costs are uniquely determined by $\theta$, and, moreover, if the matrix $\mathcal{D}_{\theta} W^{T} G(b)$ is invertible, equation (6) uniquely identifies the markup vector $\theta$. Hence, equation (6) provides an alternative to (3) to estimate costs. In Section 2.4 we study conditions for the invertibility of this matrix and for identification. 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 (5).

Note that the bidder's optimization problem in a CA is not necessarily concave. Hence, the first-order conditions (5) are not sufficient for optimality for the reduced optimization problem (4). Despite that, it is in principle possible to test computationally whether the observed bid vector that satisfies (5) is locally or globally optimal for optimization problem (4). We provide more details in the context of our application.

Similarly to equation (3), the right-hand side of equation (6) can be estimated purely from observed bidding data when it is evaluated at the observed bid vector $b$. Basically, the winning probability vector $G(b)$
and its Jacobian matrix $\mathcal{D}_{b} G(b)$ in the full-dimension model are 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 variable vector $\theta$. Equation (6) provides substantial computational advantages to the estimation process compared to equation (3). First, derivatives are now taken with respect to $d \ll A$ variables, effectively reducing the dimension of the problem. Second, instead of estimating the winning probability $G(b)$ and its Jacobian matrix, it is enough to estimate $W^{T} G(b)$ and its Jacobian matrix $\mathcal{D}_{\theta} W^{T} G(b)$. Later we show examples that illustrate how $W^{T} G(b)$ aggregates probabilities over many packages. Hence, there are much fewer probabilities to estimate, and each one has a larger value so they are easier to estimate than winning probabilities of individual packages. This makes the estimation tractable.

One apparent limitation of Assumption 5 is that the markup is additive as oppose to multiplicative to costs, which may be more appropriate. 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-characteristics. This approach is effective when the cost heterogeneity across packages can be captured by a reduced set of known variables.

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, we could choose $d=1$ so that all packages share the same markup, reducing the problem to a single decision variable. Between these two extremes there are many possible specifications for $W$. Note that data can provide guidance on what is a reasonable specification. For example, the solution of the full-dimensional first-order conditions (3) would provide information on what packages have similar markups and on what package characteristics affect markups the most. However, as we mentioned, solving (3) is intractable. In the next section we describe an alternative approach that uses the data in a tractable way to determine a reasonable specification for $W$ that balances computational efficiency versus flexibility in the markup structure.

### 2.3 Specifying the Package-Characteristic Matrix $W$

Recall that our main objective is to identify how much of the discounts observed in the bids are due to cost synergies as opposed to strategic markup adjustments when bidding for larger packages. In order for the characteristic-based markup approach to capture this type of strategic markup adjustments, it is important to allow the markup to vary in the size of a package. For this reason, we introduce a package-characteristic matrix $W$ that includes variables related to the size of each package. Doing so helps separating what portion of the volume discounts observed in the bid data arises from markup adjustments vis-à-vis cost synergies.

We describe how to incorporate the size of a package as a characteristic in $W$. Let $\left\{\mathcal{A}_{s}\right\}_{s=1}^{S}$ form a partition of the set of possible combinations $\mathcal{A}$ that groups combinations in terms of some size measure. For example, this partition could be specified in terms of the number of units in a package. Hence, $\theta_{s}$ represents the markup charged for any package of size $s$, and so the bidder needs to choose $S$ different markups, one for each possible size. The package-characteristic matrix $W \in \Re^{A \times S}$ can be specified by the indicator variables $W_{a s}=\mathbf{1}$ [package $a$ has size $\left.s\right]$. With this specification, the term $W^{T} G(b)$ in equation (6) has the
following form:

$$
W^{T} G(b)=\left[\begin{array}{c}
W_{1}^{T} G(b) \\
W_{2}^{T} G(b) \\
\vdots \\
W_{S}^{T} G(b)
\end{array}\right]=\left[\begin{array}{c}
\text { Probability of winning any package of size } 1 \\
\text { Probability of winning any package of size } 2 \\
\vdots \\
\text { Probability of winning any package of size } S
\end{array}\right]
$$

The previous size-based markup model significantly reduces the dimensionality of the problem making the estimation feasible. In particular, while 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 of the same size is a sum of these individual probabilities over a potentially large set of packages, and may be much larger. This makes the computation of the right-hand side of the first-order condition (6) tractable.

On the other hand, the size-based markup model may be too restrictive. It may be the case that two packages of the same size would have significantly different markups in the full-dimension model and should not be grouped together in the size-based markup model. We now study how these restrictions on setting markups may affect the estimation. To do this, we provide an analytical comparison of the markups estimated by the full-dimension model with those estimated via the characteristic-based markup approach. We use these results to develop a heuristic which uses the data to construct a reasonable package-characteristic matrix $W$ that refines the size-based markup model, balancing flexibility in the markup structure with computational efficiency in the estimation.

Consider a situation in which the markups of $K$ packages are aggregated into a single common markup. More formally, in the full-dimension model we have $b_{a}=c_{a}+\theta_{a}, a=1, \ldots, K$ and in this specific characteristic-based model we have $b_{a}=c_{a}+\theta_{u}, a=1, \ldots, K$, where $\theta_{u}$ is the common markup. We consider the perspective of a specific bidder and show the following proposition.

Proposition 2. Consider a given bidder placing a bid vector $b$ in a $C A$ with $K$ packages. Suppose the bids on the $K$ packages have positive winning probabilities and let $\theta_{a}, a=1, \ldots, K$ be the solution of the first-order condition of the full-dimensional model, (3). Let $\theta_{u}$ be the common markup for all K packages that solves the first-order conditions of the characteristic-based model, (6). Then,

$$
\begin{aligned}
\theta_{u} & =\frac{1}{\sum_{a=1}^{K} \alpha_{a}} \sum_{a=1}^{K} \alpha_{a} \theta_{a}, \text { where } \\
\alpha_{a} & =\frac{\partial G_{a}(b)}{\partial \theta_{u}}=\sum_{s=1}^{K} \frac{\partial G_{a}(b)}{\partial \theta_{s}} .
\end{aligned}
$$

Moreover, $\alpha_{a} \leq 0, \forall a$, with at least one $\alpha_{a}<0$.
Note that because $\alpha_{a}$ 's are negative, the common markup $\theta_{u}$ is a weighted average of the individual markups $\theta_{a}$ 's. From this we learn that, if we (the researchers) use the uniform markup model while the bidder actually solves the full-dimension model, the identified markup will be a weighted average of the individual markups. Hence, if groups of packages have similar markups in the full-dimension model, using these
groups as characteristics in the $W$ matrix would make the reduced markup model a good approximation. ${ }^{1}$
Proposition 2 shows the impact on the estimated markup of those packages that are grouped. It is also important to analyze how the grouping affects other packages not contained in the group. We begin considering a two unit case where we impose a common markup, $\theta_{u}$, for the stand-alone bids containing a single unit. The idea is to study how this grouping affects the estimated markups of the package of two units. The following proposition summarizes this result.

Proposition 3. Consider a CA with 2 units and a given bidder. Suppose that all the bids have positive winning probabilities and that $\left(\theta_{1}, \theta_{2}, \theta_{12}\right)$ solves the first-order conditions of the full-dimensional model, (3). Suppose $\left(\theta_{u}, \theta_{v}\right)$ is the solution of the first-order conditions of the characteristic-based model, (6), where $\theta_{u}$ is the common markup for single unit bids and $\theta_{v}$ is the markup for the package bid. Then,

$$
\begin{aligned}
\theta_{u} & =\lambda \theta_{1}+(1-\lambda) \theta_{2}, \\
\theta_{v} & =\theta_{12}+\gamma\left(\theta_{1}-\theta_{2}\right),
\end{aligned}
$$

where

$$
\begin{aligned}
\lambda & =\left\{\frac{\partial G_{12}}{\partial \theta_{12}}\left(\frac{\partial G_{1}}{\partial \theta_{1}}+\frac{\partial G_{1}}{\partial \theta_{2}}\right)-\frac{\partial G_{1}}{\partial \theta_{12}}\left(\frac{\partial G_{12}}{\partial \theta_{1}}+\frac{\partial G_{12}}{\partial \theta_{2}}\right)\right\} / d e t, \\
\gamma & =\left\{\left(\frac{\partial G_{2}}{\partial \theta_{1}}+\frac{\partial G_{2}}{\partial \theta_{2}}\right) \frac{\partial G_{1}}{\partial \theta_{12}}-\left(\frac{\partial G_{1}}{\partial \theta_{1}}+\frac{\partial G_{1}}{\partial \theta_{2}}\right) \frac{\partial G_{2}}{\partial \theta_{12}}\right\} / d e t, \\
d e t & =\left(\frac{\partial G_{1}}{\partial \theta_{1}}+\frac{\partial G_{2}}{\partial \theta_{1}}+\frac{\partial G_{1}}{\partial \theta_{2}}+\frac{\partial G_{2}}{\partial \theta_{2}}\right) \frac{\partial G_{12}}{\partial \theta_{12}}-\left(\frac{\partial G_{12}}{\partial \theta_{1}}+\frac{\partial G_{12}}{\partial \theta_{2}}\right)\left(\frac{\partial G_{1}}{\partial \theta_{12}}+\frac{\partial G_{2}}{\partial \theta_{12}}\right) .
\end{aligned}
$$

From the above proposition we again observe that $\theta_{u}$ is a weighted average of the individual markups. Moreover, we also observe that grouping the unit markups affects the estimated markup of the package. However, as seen in the above equation, the impact of grouping on the bundle markup, the one that is not grouped, depends on the coefficient $\gamma$ and the difference of the grouped markups. If the unit markups are very close to each other, the effect of grouping will be negligible. Quantifying the value of $\gamma$ analytically is a challenging task due to the limited knowledge on the winning probabilities. Nevertheless, computationally we have observed that in practice $\gamma$ has low values. Therefore, we expect that the effect of grouping markups on out-of-group markups will be small. In fact, our 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 not in this group.

Our previous discussion and Propositions 2 and 3 provide some guidance on how to construct a reasonable package-characteristic matrix. First, we only consider packages of the same size to form a group with a common markup. Second, we try to group markups so that the packages in the same group would have similar markups in the full-dimension model. Packages that would have significantly different markups compared to the rest of the group in the full-dimension model should be separated and have their own markup. Ideally, one could use the first-order conditions of the full-dimension model (3) to identify such

[^1]packages, but this is intractable. Instead we use the packages' winning probabilities as proxies for markup values. In fact, while this proxy is not perfect, numerical experiments suggest that packages with larger winning probabilities are likely to have larger markups in the full-dimension model. Finally, we require a minimum threshold on the aggregated winning probabilities of each package group. Recall that to estimate the markup using equation (6), we need to estimate the aggregated winning probabilities and its derivatives numerically. The minimum probability threshold is applied to ensure accurate estimation of those terms.

Based on these ideas we develop a heuristic to build the package-characteristic matrix for a given firm. The method roughly consists of three steps:

1. First, we run a simulation to estimate winning probabilities of each package; this simulation is quicker to run than solving for the first-order conditions. We identify packages that have very high winning probabilities relative to the rest. These packages are likely to have larger markups in the fulldimension model, so each of them is associated with its own markup variable. For the rest of the packages, we form several groups of packages so that packages with similar winning probabilities are grouped together. In addition, recall that we only group together packages of the same size. In short, in this step, we try to have as many markup groups as possible to the extent that computational tractability is maintained.
2. In the second step, given the candidate matrix $W$ (or equivalently, the package groups) constructed from the first step, we obtain rough estimates for the markups using equation (6). Given those estimates, we combine some of the groups together if they have similar markup levels. As suggested in the propositions of the previous subsection, this will increase computational tractability without sacrificing the flexibility of the markup structure too much.
3. This gives us the final choice of the package-characteristic matrix $W$, with which we obtain precise and final estimates of the markup vector $\theta$ through equation (6).

We use this heuristic to build the $W$ matrix for each bidder in our estimation method.

### 2.4 Identification

A condition to uniquely identify the markup vector $\theta$, and hence the costs, is that the matrix $\mathcal{D}_{\theta} W^{T} G(b)$ is invertible in equation (6). We finish this section discussing issues related to identification which are important for the specification of $W$.

In empirical settings, including the one analyzed in this paper, bidders may not submit bids on all packages. ${ }^{2}$ These unobserved package bids can still be incorporated into our framework (which assumes bids for all packages) by treating them as observed bids with very high prices that have no chances of winning. We refer to the bids which never win as irrelevant bids. In addition, some bids that are actually submitted may also be irrelevant, in the sense that they have zero probabilities of winning. For example,

[^2]this could arise as a strategic decision in order not to win a specific package when the auction rules require submission of bid prices on all packages. In what follows, we show how irrelevant bids can limit the identification of costs and how they can be handled in an actual application. We also provide necessary and sufficient conditions so that (6) identifies markups, and thereby costs.

Recall that each column of package characteristics in $W$ is associated with a markup variable in the bidder's decision $\theta$. We say that a package $a$ is associated with the markup variable $\theta_{i}$ if $W_{a i} \neq 0$, that is, the bid price of $a$ depends on the value of $\theta_{i}$. The following lemma is useful to characterize the conditions needed for identification:

Lemma 1. Consider a given bidder and auction. For any package $a \in \mathcal{A}, G_{a}(b)=0$ implies $\frac{\partial}{\partial \theta_{i}} G_{a}(W \theta+$ $c)=0$, for all $i=1, \ldots, d$.

The lemma implies that if all the bids associated with a markup variable $\theta_{i}$ are irrelevant, then the $i^{t h}$ row of $\mathcal{D}_{\theta} W^{T} G(b)$ matrix will be all zero, and the matrix will not be invertible. Because equation (6) requires invertibility of the Jacobian, the markup vector of that bidder is not identified. This problem, however, can be resolved by eliminating irrelevant bids from the model. By doing so, we can still identify markup variables as long as they have at least some relevant bids that are associated with it. We examine this issue in more detail in what follows.

Consider a given firm. Without loss of generality, we assume packages are ordered such that all the relevant bid packages (superscripted by $R$ ) are followed by the group of irrelevant bid packages (superscripted by $I$ ), so that:

$$
W=\left[\begin{array}{c}
W^{R} \\
\cdots \\
W^{I}
\end{array}\right], \quad c=\left[\begin{array}{c}
c^{R} \\
\cdots \\
c^{I}
\end{array}\right], \quad b=\left[\begin{array}{c}
b^{R} \\
\ldots \\
b^{I}
\end{array}\right], \text { and } \quad G(b)=\left[\begin{array}{c}
G^{R}(b) \\
\ldots \\
G^{I}(b)
\end{array}\right]
$$

Replacing in equation (6) we obtain:

$$
\begin{align*}
\theta & =-\left\{\left[\mathcal{D}_{\theta}\left(\left(W^{R}\right)^{T} G^{R}(b)+\left(W^{I}\right)^{T} G^{I}(b)\right)\right]^{T}\right\}^{-1}\left(\left(W^{R}\right)^{T} G^{R}(b)+\left(W^{I}\right)^{T} G^{I}(b)\right) \\
& =-\left\{\left[\mathcal{D}_{\theta}\left(W^{R}\right)^{T} G^{R}(b)\right]^{T}\right\}^{-1}\left(W^{R}\right)^{T} G^{R}(b) \\
& =-\left\{\left[\mathcal{D}_{\theta}\left(W^{R}\right)^{T} G^{R}\left(b^{R}\right)\right]^{T}\right\}^{-1}\left(W^{R}\right)^{T} G^{R}\left(b^{R}\right), \tag{7}
\end{align*}
$$

where the second to last equation follows from $G^{I}(b)=0$ and Lemma 1. In the last equation, it is implicitly assumed that the bidder only submit bids for relevant bids. Because irrelevant bids never win and by Lemma 1 small changes in the markup vector will not turn them into relevant bids, it is the same as if the bidder would not have submitted them (recall that non submitted bids are also irrelevant). Therefore, the right-hand side of equations (6) and (7) are equivalent. Consequently, the elimination of irrelevant bids will not affect the identification of the markup vector $\theta$ as long as the Jacobian in equation (7) is invertible.

Like the size-based markup model and its refinement described in the previous section, when the markup of each package is determined by one and only one markup parameter, we call it a group markup model. The following theorem provides necessary and sufficient conditions to ensure identification of the markup vector $\theta$ for the class of group markup models.

Theorem 1. Consider a given bidder and auction. Assume that the package-characteristic matrix $W$ is the specification of a group markup model. If the Jacobian matrix $\mathcal{D}_{\theta} W^{T} G(b)$ evaluated at the observed bid vector $b$ is invertible, then every markup variable has at least one relevant bid associated with it. The latter condition becomes sufficient for the invertibility of the Jacobian matrix if the following additional conditions hold: (i) the observed bid vector satisfies equation (5); (ii) the observed bid vector is such that $b-c \geq 0$; and iii) all elements of $W$ are nonnegative ( $W \geq 0$ ). In this case, the markup vector $\theta$ is uniquely identified by equation (6). ${ }^{3}$

Note that the assumption $b-c \geq 0$ is a mild rationality assumption on bidders' behavior that guarantees bidders make positive profits on each package conditional on winning that package. Also note that under the previous assumption, assuming $W \geq 0$ is essentially done without loss of generality, in the sense that for a $W$ matrix with negative entries, one can show there is another $W$ matrix with non-negative entries that produces the same markup estimates. A practical implication of the theorem is that when implementing the heuristic described in Section 2.3 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.

It is important to note that, for a given bidder, our approach only allows us to identify the cost structure of packages associated with relevant bids, that is, $c^{R}=b^{R}-W^{R} \theta$. In fact, irrelevant bids provide no information to the first-order conditions. Although it is not possible to point identify the costs of irrelevant bids, it can be shown that bounds on the costs of such "irrelevant" bid packages can be obtained. CP show that by finding the threshold bid price over which the bid becomes relevant, we can identify a lower bound on the cost of the specific irrelevant bid package. However, in a large-scale CA, this is not viable because of the computational burden. Instead, we infer the costs of those irrelevant bids using extrapolation. We will come back to this point in Section 4 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. Without this requirement, it may not be possible to point identify costs, as we illustrate with the following example. Consider a CA with 2 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).

## 3 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 provide a detailed description of the auction as well as of the data available.

[^3]
### 3.1 Brief History

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 percent 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, firstprice 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 (US\$ 577 million were awarded in 2008), being one of the largest state auctions in Chile.

For the purpose of the auction, Chile is divided into approximately 100 school districts or territorial units (TUs) in 13 geographic regions. JUNAEB holds auctions in one-third of the country every year, for around 30-35 TUs each time, awarding three-year contracts. Approximately 20 firms participate in each auction. Firms can submit bids on various groupings of TUs defining the combinatorial character of this auction. This mechanism is motivated by the belief that firms are subject to cost synergies that arise from operational advantages when serving multiple TUs. More specifically, suppliers may face economies of scale (generated by volume discounts in their input purchases) and economies of density (arising from common logistics infrastructure used to supply nearby units).

### 3.2 Auction Process

The auction process begins when JUNAEB invites and registers potential vendors. The agency then evaluates the companies from a managerial, technical and financial point of view, and eliminates those that do not meet minimum reliability standards. Qualifying vendors are classified according to two characteristics: their financial capacity (based on data from the firms' balance sheets), and their managerial competence. Usually, firms below a minimum level of managerial competence are not allowed to participate in the auction. Meal plans are standardized and service quality requirements are presented in detail. With that, firms compete on price basis. Potential vendors submit their bids simultaneously and in a single-round through an online system. Upon winning a contract, the firm receives its bid as a payment and it is responsible for managing the entire supply chain associated with all meal services in the corresponding TUs. This includes from sourcing food inputs going all the way to cooking and serving the meals in the schools.

Bidding language. A bid can cover any combination from one to eight TUs and specifies the price for which the firm would serve all meals included in the TUs in the combination. 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.

Winner determination. The allocation is chosen by selecting the combination of bids that supply all of the TUs at a minimum cost. The problem is formulated as an integer program (IP) that incorporates other considerations and side constraints. There are four types of constraints implemented in the auction and the
details of those constraints are as follows. The mathematical formulation of the IP is provided in Online Appendix A.

1. Cover all TUs: the final allocation should cover all the TUs auctioned.
2. Maximum Number of TUs: There is a maximum number of TUs that each firm can be allocated in any given auction. This maximum is based on the financial evaluation conducted by JUNAEB every year and therefore can be different across firms and auctions, ranging from 1 to 8 TUs.
3. Global Market Share Constraints: To avoid excessive concentration and encourage diversification, at any point in time, the total standing contracts of any firm cannot exceed $16 \%$ of the total number of meals included in all TUs in the entire country. Hence, depending on the volume of standing contracts, the maximum volume can be also different across firms and auctions.
4. Local Constraints: To facilitate supervision and control of the firms, there are constraints on the maximum number of firms serving in each geographical region. On the other hand, to actively respond to contingencies such as bankruptcies, there are also constraints on the minimum number of firms serving in each geographical region. Geographical regions in low population areas contain less than five TUs while regions with higher population typically contain between 10 and 20 TUs.
5. Global Competition Constraint: For similar reasons as the global market share constraints, there is a constraint in the minimum number of firms winning contracts in each auction (this number can vary across auctions, but is around 10).

Finally, we note that in the structural method described in Section 2.1 we assumed that the auction allocates at most one package per firm. While this restriction is not imposed in our empirical application, except for isolated exceptions, firms actually win at most one package in practice. Hence, this is a reasonable assumption even in our application, so we also impose the one-package per-firm constraint in the winner determination problem.

### 3.3 Description of the Data

We collected data on all auctions between 1999 and 2005. Our data set contains all bids placed by all firms in each auction and the identity of winning firms in each auction. In addition, we also have detailed information on the auction parameters and characteristics of all participating firms. A more detailed description of the data can be found in Olivares et al. (2011).

We have detailed information on the parameters of each auction, including the TUs auctioned. For each TU, we know its annual demand in terms of number of meals to be served and the geographic location of its schools. TUs are heterogeneous in terms of size and the density of the school population, which are key factors affecting firms' supplying costs. In addition, we have all the parameters used in the constraints in the integer program associated with the winner determination problem.

For each auction, we know the identity of all participating firms and their characteristics as well as all the bids placed by them. In addition, we know the firm limits for the maximum number of TUs and market
share constraints in each auction. Finally, we 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.

## 4 Estimation

This section describes details of the estimation of the structural model developed in Section 2 using the data from the Chilean School Meals auction that was described in Section 3. Our structural model identifies the costs of each firm based on equation (6). Similarly to CP, we use a two-stage method. In the first stage, described in Section 4.1, we parametrically estimate the distribution of competitors' bids, $H(\cdot)$, from the data. In the second stage, we estimate the winning probabilities of the package bids, $W^{T} G(b)$, and its Jacobian using simulation (see Section 4.2). We replace these values in equation (6) to obtain point estimates of the markups, and thereby the cost of each package. Section 4.3 provides the estimation results.

### 4.1 Estimating the Distribution of Competitors' Bids

As mentioned in Section 2.1, we conduct an auction-by-auction estimation of the distribution of competitors' bids. Given the high dimensionality of the bid vector $b_{f}$, which in our application is in the order of hundreds of packages, it is unfeasible to estimate the multivariate distribution of this bid vectors non-parametrically as was done by Guerre et al. (2000) in the single-unit auction case. Even in a small CA, Cantillon and Pesendorfer (2006) use a parametric approach. They assume a log-normal multivariate distribution, imposing constraints in the variance-covariance matrix. We follow a similar approach, as discussed next.

While the parametric approach provides a feasible estimation method in large-scale CAs, it is important to incorporate the needed flexibility so that the relevant factors that affect bid prices are captured in the model. First, in our application there is significant heterogeneity regarding the costs of serving different territorial units. These differences arise primarily because of location and density of schools; for example, units located in isolated rural areas tend to be more expensive than units in urban areas. Moreover, there may be substantial heterogeneity across firms' costs. For example, some firms have national presence, are vertically integrated, and have well functioning and efficient supply chains; while others are more rustic local firms. Hence, it is important to allow for sufficient flexibility in the distribution of bids to incorporate these two types of heterogeneity.

Second, package bids of the same bidder may be correlated. In our application, 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 in the provision of services, a supplier with a low cost for a unit may also have low costs in 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. We note that, as discussed in previous work by Olivares et al. (2011), the correlation structure of the competitors' bids has direct implications on the incentives to engage on strategic markup adjustments. Consequently, allowing for a flexible covariance structure that captures differences in the correlations between units can be important to identify the markups chosen by the different firms.

Third, bid data exhibits significant discounts both due to scale and density. Motivated by Olivares et al. (2011), we develop a parametric econometric model for package bids which captures these discounts, unit/firm heterogeneity and correlation between bids. Let $v_{i}$ denote the volume of unit $i$, measured in million meals per year, and $v_{a}=\sum_{i \in a} v_{i}$ the total volume of package $a$. Bid prices are modeled as follows: ${ }^{4}$

$$
\begin{equation*}
b_{a f}=-g^{s c a l e}\left(v_{a}, \beta_{s(f)}^{s c a l e}\right)-\sum_{c \in C l(a)} g^{\text {density }}\left(v_{c}, \beta_{s(f)}^{\text {density }}\right) \cdot \frac{v_{c}}{v_{a}}+\sum_{i \in a} \tilde{\delta}_{i f} \frac{v_{i}}{v_{a}}+\tilde{\varepsilon}_{a f} . \tag{8}
\end{equation*}
$$

The dependent variable, $b_{a f}$, with some abuse of notation, denotes the per-meal price 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 (8) capture: (1) the effect of discounts due to scale ( $g^{\text {scale }}$ ); (2) the effect of discounts due to density ( $g^{\text {density }}$ ); (3) the effect of the specific units contained in the package; and (4) a Gaussian error term $\tilde{\varepsilon}_{a f}$ capturing other factors affecting the bid price.

The distributional assumptions we make on the variables in equation (8) will induce distributions of competitors' bids $H\left(\cdot \mid Z_{f^{\prime}}\right)$ for the auction game of asymmetric information played by firms (see Assumption 4). An important distinction is which parameters of this equation are assumed to be known to the bidder at the time of choosing its bid, and which ones are unknown and thereby considered random from the bidders' perspective. These random parameters capture the asymmetric information among the bidders. This distinction between known and unknown parameters in equation (8) is important for simulating winning probabilities.

We use tilde (e.g. $\tilde{\delta}_{i f}$ ) to denote factors that are unknown and random to the bidder. Hence, from the perspective of a bidder, model (8) is a regression with error components determined by $\left\{\tilde{\delta}_{i f}\right\}_{i \in A}$ and $\left\{\tilde{\epsilon}_{a f}\right\}_{a \in \mathcal{A}}$ and all the asymmetric information among bidders is encapsulated in those random components. In the context of the auction game, the parameters characterizing the discounts ( $\beta^{\text {scale }}, \beta^{\text {density }}$ ) and the distribution of the error components $\left\{\tilde{\delta}_{i f}\right\}$ and $\tilde{\epsilon}_{a f}$ are considered common knowledge (as well as the bid data generating process specified by equation (8)). Therefore, to determine the distribution of competitors' bids as seen by the bidder, we need to estimate ( $\beta^{\text {scale }}, \beta^{\text {density }}$ ) and the distribution of $\left\{\tilde{\delta}_{i f}\right\}$ and $\tilde{\epsilon}_{a f}$. Next, we discuss details of this estimation.

First, consider the terms capturing scale and density discounts, $\left(\beta_{s(f)}^{s c a l e}, \beta_{s(f)}^{\text {density }}\right)$. The model allows for some observed heterogeneity of these discounts across firms, with $s(f)$ indicating groups of firms of different business sizes. We found significant differences in discounts across the largest firms and the rest of the firms, so we grouped bidders into two groups, $s \in\{L, O\}$ (for Large and Other), to estimate discounts ( $L$ firms refer to the largest firms in JUNAEB's classification and can bid on packages of up to 8 units). There are several reasons that can help explain differences in discounts among firms of different size. First, bigger firms operate at a different scale and tend to operate other businesses outside the school meals procurement system. Hence, synergies for these firms could be different, which in turn could lead to different discounts. Note, however, that the discount functions $g^{\text {scale }}$ and $g^{\text {density }}$ should not be interpreted directly as cost synergies because part of the discounts could arise from strategic behavior. In this regard, strategic markup adjustments could be different for firms that can bid on bigger packages, leading to further differences in the

[^4]discounts. We assume that the parameters $\left\{\beta_{s(f)}^{\text {scale }}, \beta_{s(f)}^{\text {density }}\right\}_{s(f) \in\{L, O\}}$ are common knowledge and that all the uncertainty associated to the magnitude of the discounts is provided by the error terms $\tilde{\epsilon}_{a f} .^{5}$

To measure scale discounts per meal, $g^{s c a l e}$ is specified as a step function of the package size $v_{a}$, and therefore the total discount in a package is a piece-wise linear function of the package size $v_{a}$. In contrast, density discounts depend on the proximity of the units in the package. To capture this, $g^{\text {density }}$ depends on the size of clusters of units in a package, where a cluster is a subset of the units in package $a$ which are located in close proximity. In equation (8), $C l(a)$ denotes the set of clusters in the package and $c$ indicates a given cluster in this set, with size $v_{c}$. This approach follows directly from the work of Olivares et al. (2011) and further details on the computation of clusters are described in the appendix of that article.

Consider now the next-to-last term of the right-hand side of equation (8); this error-component term is a weighted average of the parameters $\tilde{\delta}_{i f}$, which are firm and unit specific, capturing the effects of the individual units contained in package $a$. The $\tilde{\delta}_{i f}$ 's can be viewed as an individual price that bidder $f$ is implicitly charging for unit $i$, net of any scale and density discounts. Note that this needs not be equal to the price the bidder charges for the stand-alone bid for unit $i$, because $\tilde{\delta}_{i f}$ is an average implicit price considering all packages that contain the unit. These implicit prices could vary with the unit characteristics (e.g., urban vs. rural units) and local advantages of a firm in that unit, among other factors.

Note that we assume that the bidder does not observe the $\tilde{\delta}_{i f}$ 's of the competitors, which adds uncertainty to the competitor's bid distribution as observed by the bidder when choosing its bid. Instead, bidders view each competitor's $\tilde{\delta}_{f}=\left\{\tilde{\delta}_{i f}\right\}_{i \in U}$ as a random vector drawn from a distribution which is common knowledge. Hence, we are not interested in point estimates of $\tilde{\delta}_{f}$ 's per-se but rather the distribution of these average implicit prices as perceived by bidders. Accordingly, we let the vector of average implicit prices $\tilde{\delta}_{f}$ follow a multi-variate normal distribution and we seek to estimate the mean and covariance matrix of this distribution. Since our application has about 30 units on each auction, we need to impose further restrictions to estimate the covariance matrix. The specification we propose captures two important elements that are important for this application. First, some of the observed firms' characteristics, denoted by the vector $Z_{i f}$, affect the implicit prices charged for the units. For example, Olivares et al. (2011) show that firms that seek to renew a contract they are already serving tend to offer more competitive prices. These firm characteristics are observed by all bidders and therefore considered common knowledge. Second, there is spatial correlation among the units, so prices for a unit tend to be positively correlated with the prices of nearby units. The following specification captures both effects:

$$
\begin{equation*}
\tilde{\delta}_{i f}=\bar{\delta}_{i}+\beta^{Z} Z_{i f}+\psi_{r(i), f}+\nu_{i f} \tag{9}
\end{equation*}
$$

The parameters $\left\{\bar{\delta}_{i}\right\}_{i \in U}$ are treated as fixed-effects which capture the average implicit price for each unit charged among all bidders. $Z_{i f}$ denote the aforementioned unit/firm characteristics. The term $\psi_{r(i), f}$ is an error component associated with a pre-specified geographic region $r(i)$ where unit $i$ is located; there are $R$ different pre-specified regions. The error components $\left(\psi_{1 f}, \cdots, \psi_{R f}\right)$ follow a multi-variate normal distribution with zero mean and covariance matrix $\Omega$. Finally, the remaining error term $\nu_{i f}$ follows an independent zero-mean normal distribution and is heteroskedastic with variance $\sigma_{i}^{2}$. These distributions are

[^5]common knowledge. Because $R$ may be much smaller than the number of units, this specification provides a substantial dimensionality reduction over the fully flexible distribution of $\tilde{\delta}_{i f}$ 's.

Under the specification (9), the covariance structure of any two average implicit prices $\tilde{\delta}_{i f}$ and $\tilde{\delta}_{j f}$ is given by: $\operatorname{Cov}\left(\tilde{\delta}_{i f}, \tilde{\delta}_{j f}\right)=\Omega_{r(i), r(j)}+\sigma_{i} \sigma_{j} \mathbf{1}[i=j]$. 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 pattern is observed in the data. However, it is flexible in allowing positive or negative correlation among units in different regions.

In summary, the competitors' bid distribution $H\left(\cdot \mid Z_{f} ; \phi\right)$, where $\phi$ is the set of distribution parameters, is a mixture distribution described by the following terms: (1) the deterministic component associated with the discounts $g^{\text {scale }}$ and $g^{\text {density }}$, captured by the vector parameters $\beta^{\text {scale }}$ and $\beta^{\text {density }}$; (2) a random component associated with the average implicit price vector $\tilde{\delta}_{f}$, whose distribution is fully described by the vector parameters $\bar{\delta}=\left(\bar{\delta}_{1}, \cdots, \bar{\delta}_{N}\right), \beta^{Z}, \sigma=\left(\sigma_{1}, \cdots, \sigma_{N}\right)$ and the covariance matrix $\Omega$; and (3) a Gaussian error component $\tilde{\varepsilon}_{a f}$, which we assume to be heteroskedastic so that the variance depends on the number of units in the package, $\sigma_{|a|}^{\varepsilon}$. All of these parameters that characterize the distribution of competitors bids are considered common knowledge and need to be estimated from bidding data, so that we can use the competitors' bid model in the simulation of the winning probabilities.

The following two-step procedure is used to estimate the econometric model defined by equations (8) and (9):

- First step: estimate (8) via a Generalized Least Squares (GLS) regression to obtain estimates of $\beta^{\text {scale }}$, $\beta^{\text {density }}$, and point estimates of the implicit prices $\tilde{\delta}_{i f}$ 's.
- Second step: plug-in the estimated $\tilde{\delta}_{i f}$ 's into equation (9) and estimate its model parameters through maximum likelihood.

The identification of model (8) is based on variation across package bids in a single auction, and hence requires a large number of package bids. More specifically, the two step procedure described above estimates scale and density discounts using variation across different combinations submitted by the same firm over the same set of units. Under the usual orthogonality conditions of GLS, the first step regression provides consistent estimates of $\beta^{\text {scale }}, \beta^{\text {density }}$ and point estimates of the $\tilde{\delta}_{i f}$ 's. ${ }^{6}$

Identification of the parameters in model (9) is based on variation across units and firms. Given consistent estimates of the implicit unit prices $\tilde{\delta}_{i f}$, the second step provides consistent estimates of $\left\{\bar{\delta}_{i}, \sigma_{i}\right\}_{i \in U}, \beta^{Z}$ and $\Omega$ as long as $Z_{i f}$ is orthogonal to the error components $\psi_{r(i), f}$ and $\nu_{i f}$. The consistency of our two-step method is a special case of the 2 -step M-estimators described in Wooldridge (2002).

Estimates for the Parameters of Bid Distribution. We provide the results for the 2003 auction. Table 1 reports estimates of $\beta^{\text {scale }}$ and $\beta^{\text {density }}$ from the first step regression. The scale and density per-meal

[^6]discount curves, $g^{\text {scale }}\left(v_{a}, \beta_{s(f)}^{s c a l e}\right)$ and $g^{\text {density }}\left(v_{c}, \beta_{s(f)}^{\text {density }}\right)$, are specified as step functions with interval of three million meals per year in the package volume $v_{a}$ and cluster size $v_{c}$, respectively. Each number indicates the average discount in per-meal price when units are combined to form a package that 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 $\mathrm{Ch} \$ 22.78$ 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 $\mathrm{Ch} \$ 11.27$ on average for a large firm. The results show that large firms were able to provide higher discounts which amounts up to $9 \%$ of average bid price. All the coefficients are estimated with precision at the significance level of $0.01 \%$.

In addition, the first-step estimation of regression (8) provides point estimates of the implicit average prices $\delta_{i f}$ 's (not shown). On average, the standard error for these estimates are in the order of $0.5 \%$ of the point estimates, which is reasonably accurate. To further validate these estimates we compared them with the stand-alone bids $b_{i f}$ 's. The ratio $b_{i f} / \delta_{i f}$ is on average 0.998 with standard deviation of 0.026 . The correlation of the two measures is 0.987 . These results suggest that the implicit average prices effectively separate out the individual prices from package discounts.

The second step estimation of regression (9) provides estimates for the distribution of the average implicit prices $\delta_{i f}$ 's, characterized by $\left\{\bar{\delta}_{i}, \sigma_{i}\right\}_{i \in U}$, the covariance matrix $\Omega$ and $\beta^{Z}$, the coefficients of the firm characteristics. Firm characteristics include an indicator on whether the firm was awarded the unit in the previous auction (we also tried other firm characteristics but those were not statistically significant). Due to 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 the incumbency effect $\left(\beta^{Z}\right)$ is -6 with a p-value of 0.012 , suggesting that on average incumbent firms submit bids that are around $1.5 \%$ cheaper than non-incumbent firms.

Table 2 shows the correlations between the region effects $\psi_{r(i), f}$ (which were calculated based on estimates of the variance/covariance matrix $\Omega$ ). 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. The last column of the table shows the standard deviations of each region effect $\psi_{r(i), f}$ (which corresponds to $\sqrt{\Omega_{r r}}$ for each region $r$ ). All the standard errors of the second step estimates are computed via a parametric bootstrapping procedure.

### 4.2 Markup and Cost Estimation

Using the estimated distribution of competitors' bids, markups are estimated using equation (6) for each firm. This requires calculating the aggregated winning probabilities $W^{T} G(b)$ and its Jacobian, as described in this section.

First, the specification of the package-characteristic matrix $W$ is based on a refinement of the sizedbased markup approach described in section 2.3. Package size is measured in terms of the number of units in the combination, so that the packages with same number of units are grouped together. Each size group is then partitioned into those that should have similar markup levels in the full-dimension model, based on the heuristic described in section 2.3.

In our application, package volume - defined as $v_{a}=\sum_{i \in a} v_{i}$ - has a first-order effect on the bid price and so prices may vary substantially even within each size group. For this reason, we assume that bids within each group have a common per-meal markup, instead of a fixed absolute markup. Defining $b_{a}$ and $c_{a}$ as the per-meal bid and per-meal cost of package $a$, and defining the non-zero components of $W$ as $W_{a s}=v_{a}$, the firm's decision variable $\theta$ can be interpreted as a per-meal markup vector. ${ }^{7}$

Next, we describe a Monte Carlo simulation to calculate the winning probabilities for a given bidder $f$. Given the point estimates $\hat{\phi}$ for the distribution parameters from our two step approach, each simulation run $l$ consists of the following:

1. For each competitor $f^{\prime}$ (different from bidder $f$ ), draw independently a bid vector $b_{f^{\prime}}^{(l)}$ (containing all submitted packages by that firm) from the estimated bid distribution $H\left(\cdot \mid Z_{f^{\prime}} ; \hat{\phi}\right)$.
2. Using the observed bid $b_{f}$ for bidder $f$ and the sampled competitors' bids $\left\{b_{f^{\prime}}^{(l)}\right\}_{f^{\prime} \neq f}$, solve the winner determination problem.
3. Let $\iota^{(l)}$ be a vector of $A$ binary variables indicating the packages awarded to the bidder $f$. Store in memory the vector $W^{T}{ }_{l}{ }^{(l)}$.

At the end of the simulation after $L$ replications, the aggregated winning probability vector can be estimated by:

$$
W^{T} G(b) \approx \frac{1}{L} \sum_{l=1}^{L} W^{T} \iota^{(l)} .
$$

Note that if the distribution of competitors' bids is estimated consistently, then the previous equation provides consistent estimates of the aggregated winning probabilities as $L$ becomes large. Finite differences are used to compute the Jacobian of $W^{T} G$, which requires calculating the change in the winning probabilities from a small change in each markup variable $\theta_{i}$. Because the bid is linear in $\theta$ (by Assumption 5), this is equivalent to consider a small change in the observed bid vector $b$ in the direction of that markup variable. Specifically, consider a change in the $j^{\text {th }}$ component of the markup vector, and let $W_{i}$ be the $i^{\text {th }}$ column of the package-characteristic matrix $W$. To calculate the $i^{\text {th }}$ row and $j^{\text {th }}$ column element of the Jacobian we use the central finite difference method:

$$
\left[\mathcal{D}_{\theta} W^{T} G(b)\right]_{i j}=\frac{\partial W_{i}^{T} G(b)}{\partial \theta_{j}} \approx \frac{W_{i}^{T} G\left(b+h W_{j}\right)-W_{i}^{T} G\left(b-h W_{j}\right)}{2 h}
$$

The computation of $W_{i}^{T} G\left(b+h W_{j}\right)$ and $W_{i}^{T} G\left(b-h W_{j}\right)$ is done via simulation as before: on each simulation run, we solve the winner determination problems with the perturbed bid prices and keep track of the winning bids. Once the aggregated winning probability vector $W^{T} G(b)$ and its Jacobian matrix $\mathcal{D}_{\theta} W^{T} G(b)$ are estimated, the markup vector $\theta$ for this bidder is obtained through the identification equation (6): $\theta=-\left\{\left[\mathcal{D}_{\theta} W^{T} G(b)\right]^{T}\right\}^{-1} W^{T} G(b)$. The winning records used in the central finite difference method

[^7]also enable to estimate the direct second-order derivatives of the bidders' expected profits with respect to each of their markup variables. We obtained that these estimates are negative for all firms, which is consistent with the local optimality of the estimated markups. ${ }^{8}$

### 4.3 Estimation Results

The estimated markups reveal reasonable levels for most of firms with average margin of $2.7 \%$. However, from the 20 participating firms there are two extreme firms whose markups are unreasonably high and lead to negative costs for some packages. This indicates that the assumptions on the bidders behavior may not be satisfied for these firms. Because we are unable to correctly infer the cost information of these firms and because these firms did not win any contracts in the 2003 auction, we omit them from our analysis hereafter.

We observe roughly three groups of firms. The first group, which we call "aggressive" group, consists of nine firms whose total winning probabilities aggregated over all packages are higher than $50 \%$. The other firms have very low winning probabilities (less than $2 \%$ ) except for two firms whose total probabilities are $44 \%$ and $14 \%$. In terms of markups, the aggressive firms set markups ranging from $2 \%$ to $18 \%$ of the average bid price with an average markup of $4.7 \%$ of the average bid price (US\$ 0.88 per meal). As expected, the other firms set lower markups, resulting in an average markup overall firms of around $2.8 \%$ of the average bid price. Table 3 shows the average per-meal markup estimates of each package size for three representative firms of different level of "aggressiveness". In addition, the estimates indicate that firms reduce their markups as the size of packages increases, showing that some portion of the discounts in package bids are due to markup adjustments.

Firms submit hundreds to thousands of bids, and about $13 \%$ of them are relevant bids. However, for the aggressive group, this increases to $22 \%$. In addition, as we have 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 was US\$ 70.5 million per year and the supplying costs was US\$ 67.2 million per year, which yields $4.8 \%$ of average profit margins to winning firms. This level of profit margins is consistent with the Chilean government's estimate for this market, which is re-assuring.

Finally, as a robustness check, we also performed the estimation for the 2005 auction, where 16 firms participated for 23 units. The results look 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 give $3.5 \%$ of average profit margins to winning firms. ${ }^{9}$

[^8]Next, we evaluate the cost synergies - cost savings from combining units together - implied by the estimates. As mentioned in Section 3, there could be two important sources of synergies in our application: i) synergies due to economies of scale, which depend on the total volume of the package that is supplied; and ii) synergies due to economies of density achieved when nearby units are supplied together. Next, we describe how to calculate both type of synergies based on the estimated markups.

The per-meal cost of each package $a$ submitted by firm $f$ is given by $c_{a f}=b_{a f}-w_{a} \theta_{f} / v_{a}$, where $\theta_{f}$ is the markup vector estimated for that firm, $b_{a f}$ is again the per-meal bid price placed by firm $f$ for package $a$, and $w_{a}$ is the $a^{t h}$ row of package-characteristic matrix $W$ used for bidder $f$. The per-meal cost synergy in this package, denoted by $s_{a}$, can then be calculated as $s_{a}=\sum_{i \in a} \frac{v_{i}}{v_{a}} c_{i}-c_{a}$, where $c_{i}$ is the point estimate for the cost of unit $i$. Table 4 shows some summary statistics of the cost synergies. The results of this calculation suggest there are significant cost synergies amounting up to around $5 \%$ of the average bid price.

One disadvantage of the synergies that are shown in Table 4 is that they rely only on the sample of relevant bids, with a few firms accounting for a disproportionate fraction of this sample. This is because this direct calculation of cost synergies requires the costs of single-unit packages, and many of these single-unit bids are irrelevant. In order to use a larger portion of the packages to estimate cost synergies, we run a regression similar to (8) but where the dependent variable is the estimated per-meal cost of the package (rather than the per-meal bid). Here, we also try to disentangle how much of these cost synergies arise from economies of scale and density:

$$
\begin{equation*}
c_{a f}=\sum_{i \in a} \xi_{i f} \frac{v_{i}}{v_{a}}-g^{\text {scale }}\left(v_{a}, \gamma_{s(f)}^{s c a l e}\right)-\sum_{c \in C l(a)} g^{\text {density }}\left(v_{c}, \gamma_{s(f)}^{\text {density }}\right) \cdot \frac{v_{c}}{v_{a}}+\varepsilon_{a f}, \tag{10}
\end{equation*}
$$

where again $s(f) \in\{L, O\}$ indicates one of the two firm group sizes. As in (8), $g^{\text {scale }}$ and $g^{\text {density }}$ are specified as step functions of the size of the package $\left(v_{a}\right)$ and the cluster $\left(v_{c}\right)$, respectively. The parameters $\xi_{i f}$ represent an implicit cost of each unit, which is estimated for all units, including those units for which the single-unit package was irrelevant. Note that for this regression we use all relevant packages. To validate this approach, we compared the estimates of these implicit costs with those estimated directly via the structural estimation (over the sample of relevant single-unit bids). The correlation between the two is 0.9931 and the average absolute difference is $1.04 \%$ of average unit cost. The ratio $c_{i f} / \xi_{i f}$ averages to 1.000 with standard deviation of 0.011 . Hence, equation (10) seems a reasonable approach to estimate cost synergies.

Figure 2 shows the estimated cost synergies (from equation (10)) together with the bid discounts estimated previously from equation (8) over the sample of relevant bids. The results are shown for each of the firm groups. ${ }^{10}$ The results show that while there is some strategic markup adjustments, most bid discounts are actually explained by cost synergies. These synergies are quite significant and can be as large as $6 \%$ of the bid price on average. The results also show that economies of scale are predominant, but that economies of density are also important and can account for $1 \%$ of the average bid price.

Finally, similar to equation (8) and equation (10), regression equations without the density terms

[^9]$g^{\text {density }}$ will capture the average discount and synergy levels in terms of size of the packages. This information is also useful because it gives us the overall level of bid discounts and cost synergies due to the combined effect of economies of scale and density. Figure 3 shows the estimated overall bid discounts and cost synergies of those 16 firms using the relevant bids and the corresponding costs. The results suggest that on average, most of the bid discounts (at least $75 \%$ ) are driven by cost synergies as oppose to strategic markup adjustments.

## 5 Efficiency Analysis and Counterfactuals

The previous results seem to suggest that in our application allowing package bidding may be appropriate: cost synergies are significant and account for most bid discounts vis-á-vis strategic markup adjustments. Moreover, the overall markups that firms gain do not seem too large, resulting in a reasonable total procurement cost. In addition, the results suggest that the bidding language should allow bidders to express both economies of scale and density. 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).

While suggestive, the previous results are not conclusive. In this section we use our estimates to provide sharper results concerning the efficiency and procurement cost 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. We provide results for the 2003 auction. ${ }^{11}$

### 5.1 Performance of the First-Price CA

In this section we study the allocative efficiency of the first-price CA. The winning bidders' costs under the first-price CA allocation can be directly computed using the cost estimates obtained in Section 4.2. If we had the cost estimates for all possible unit combinations, one could also calculate the minimum-cost allocation. Unfortunately, the structural estimation only identifies the costs of relevant bids, and the minimum-cost allocation over this subset of combinations could overestimate the true minimum-cost allocation that also considers irrelevant bid combinations.

To address this issue, we propose estimating the cost of irrelevant bid packages through an out-ofsample extrapolation based on equation (10). However, the total number of feasible packages are in the order of millions and it is computationally infeasible to extrapolate to the entire set of (out-of-sample) packages. Instead, we choose the set of packages on which at least one bidder placed a bid, which is in the order of 30 thousand packages. We call this the expanded package set. Then for each firm, we extrapolate costs to this expanded package set as long as the package satisfies the maximum TU and global market share constraints for that firm. While this is a small subset of the entire packages, it provides a reasonable approach to extending the set of bids observed in the data.

[^10]This 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 minimum-cost allocation which 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 cost efficient allocation that minimizes costs over the set of relevant bids and feasible allocations (considering the constraints described in Section 3) is equal to US\$ 66.7 million per year. If we also consider the cost extrapolation to irrelevant bids as described above (i.e., over the expanded package set), the minimum-cost allocation goes down to US\$ 66.2 million per year. This implies an efficiency loss of the first-price CA of $1.5 \%$, which is evidently low. ${ }^{12}$ The high efficiency and relatively small profit margins for firms (around $5 \%$ as presented in Section 4.3) achieved by the school meals CA suggests that it is a reasonable mechanism for the procurement of this public service.

### 5.2 The VCG Mechanism

While our previous results supports that using the first-price CA in our application seems appropriate, it is also useful to compare the performance with alternative auction mechanisms. Doing these counterfactuals on alternative mechanisms requires computing the bidding strategies played in equilibrium by the bidders. Unfortunately, there are few equilibrium results for most of the multi-unit auction mechanisms that are used in practice. For example, it is intractable to compute equilibria of the respective games of asymmetric information for our CA, for independent single-unit auctions under the presence of synergies, or for the multi-round CAs used by the FCC in the wireless spectrum auctions.

We can perform, however, a counterfactual with the Vickrey-Clarke-Groves (VCG) mechanism, which is a generalization of a second price auction: the payment to a winner is essentially the cost of providing the units she wins in the lowest cost allocation without her, and loosing bidders do not receive payments. It is well known that under this payment rule, truth-telling is a dominant strategy, i.e., bidders report their true costs. In Online Appendix C we provide details about VCG and its payment rule. Like in the first price CA, winners in VCG are also determined by finding the combination bids that achieve the minimum procurement cost; this results in the efficient allocation due to the truth-telling property. However, despite these theoretical virtues, VCG mechanisms have been criticized for other numerous drawbacks, leading to a very rare use in practice. In particular, Ausubel and Milgrom (2006) have shown that in the face of complementarities, the VCG procurement costs can be prohibitively high. This and other deficiencies of VCG in settings with complementarities have motivated an active research agenda in recent years that studies alternative payment

[^11]rules, giving rise to the so-called "core-selecting auctions" (we provide more details below). Hence, it is on itself interesting to see how VCG mechanisms would perform in real-world applications.

In our analysis we use the same set of extrapolated bids as in Section 5.1 as the bids (costs) that bidders would report in a VCG mechanism. We know the VCG allocation is efficient, so it coincides with the minimum-cost allocation (that satisfies all constraints described in Section 3.2); this was previously computed in Section 5.1.From the bids, we can compute the individual VCG payments to the winning bidders (see online appendix), 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.

Given the theoretical literature mentioned above describing the pitfalls of VCG, the result is striking; in our application, VCG achieves payments comparable to the first-price CA, so in fact VCG induces reasonable procurement costs. 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 her bids are eliminated, there is another allocation that achieves costs close to the minimum-cost allocation, leading to reasonably low VCG payments. More broadly, it is interesting to note that in the examples provided by Ausubel and Milgrom (2006), the amount of competition is limited, resulting in high VCG payments. We believe that VCG should achieve reasonable procurement costs in settings with a reasonable amount of bidders that are able to submit many package bids. The latter should be expected when it is relatively effortless for a bidder to evaluate its cost in an additional package, therefore, the bidder can easily submit many package bids.

We finish this subsection by observing that Ausubel and Milgrom (2006) and Milgrom (2004) show that the poor performance of VCG arises from the fact that the VCG outcome may not be in the core of the transferable utility cooperative game played among the bidders and the buyer (auctioneer). In this sense, the core can be understood as a competitive benchmark; if the outcome is not in the core, the payments are so high that there is a group of bidders that can offer a more favorable deal to the auctioneer. In our application, we find that indeed the VCG payoffs lie essentially in the core, which is consistent with the reasonable total procurement cost achieved by VCG. In particular, following Day and Raghavan (2007) we find the closest point in the core (with respect to the truthful bids) to the VCG payments under a suitable norm. We find that the differences in total procurement costs between these two points is only $0.1 \%$ in 2003. Also, individual payments are very similar as well; half of the winners receive exactly the same payments in the core point as in VCG, and the rest receive payments that are no more than $0.7 \%$ apart. We provide more details about the core analysis in Online Appendix C.

### 5.3 Supplier Diversification

The CA of our application imposes three types of constraints aimed at preserving a more diversified supplier base: (1) a single bidder cannot be awarded more than $16 \%$ of the total volume including outstanding contracts awarded in previous years (market share constraint); (2) a minimum number of winning firms on each auction (global competition constraint); and (3) a minimum number of winning firms on each of the 13
pre-specified geographic regions (local constraints). We now focus on measuring what is the efficiency loss imposed by these constraints.

To study efficiency of the first-price CA, we have already calculated the minimum-cost allocation that satisfies these constraints. We could compare this with the minimum-cost allocation obtained under the larger feasible set of allocations when the constraints are removed. However, this may not be a fair comparison because bids on packages that violate some of the constraints are not submitted by the bidders. In other words, in the counterfactual world without the constraints we should observe new package bids that are not observed under the current format with the constraints. To address this issue, we expand the set of submitted bids in the counterfactual without the constraints as we now explain.

First, consider the market share constraint that imposes a maximum volume of $16 \%$ of the total volume of the country, equivalent to about $K=40$ million meals per year to each firm. Under this constraint, bids on packages with larger volume than $K$ will never be observed in the data. It turns out that because of the 8 unit limit for the packages, the market share constraint is never binding for those firms which do not have any existing outstanding contracts, because the maximum volume that can be achieved with 8 units is less than $K$. So those firms do place bids on packages of any volume with at most 8 units. We call such firms whose bidding is not limited by the market share constraint the unrestricted firms.

To extrapolate costs to packages violating the market share constraint we do the following. Consider a large bidder $f$ that has existing outstanding contracts for a total volume of $X$. This firm can only submit packages of volume less than or equal to $K-X$. Removing the constraint would allow this bidder to submit packages of any volume up to $K$, as long as they have 8 units or less. We denote by $A_{X}$ the set of observed combinations that are infeasible for bidder $f$ but feasible for the unrestricted bidders, hence contained in the expanded package set. We can use regression (10) to predict the costs of combinations in $A_{X}$ for bidder $f$. Doing this for all bidders allows us to build a larger feasible set that contains bids that would not be feasible when the $16 \%$ market share constraint is included. Again the expanded package set which is in the order of 30 thousand packages - still less than the 20 million possible packages that could be submitted - and provides a reasonable set of bids to evaluate the effect of removing the market share constraints.

In contrast, packages that violate the local competition constraints are almost never observed. To illustrate why this is the case, consider region 13 which has seven units but the minimum number of firms required to win is four. Hence bids on any package containing five of more units in region 13 will violate this constraint and will never win. Note that unlike the market share constraint which is a firm-wise restriction, local constraints are applied to all firms and hence no such packages are found in the expanded package set. For this reason, we cannot analyze the effect of removing the local constraints using the same approach to expand the set of bids. Finally, we note that it is not a priori clear whether removing the global competition constraint would result in significantly different package bids submitted, because in any event firms cannot submit packages larger than 8 units. Therefore, we do not include additional bids associated to removing that constraint, and we focus on the efficiency loss caused by the market share constraint and the global competition constraint.

To measure the efficiency loss due to the market share and the global competition constraints, we compare the minimum-cost allocations with all those constraints and without the two types of constraints.

We find that this efficiency loss in 2003 is about $0.57 \%$, which is relatively small. ${ }^{13}$ The small impact of these constraints on efficiency can be partially explained by the structure of the cost synergies in the industry. As we saw in section 4.2 , scale cost synergies get exhausted, so there are small cost reductions for combinations that lie beyond the volume range that is currently feasible in the auction. To further evaluate the inclusion of these constraints in the mechanism, it would be useful to measure the value that the constraints aimed at promoting supplier diversification bring in terms of increased competition. We leave this analysis for future research. ${ }^{14}$

## 6 Conclusions

In this paper we develop a structural estimation approach for large-scale first-price CAs and applied it to the Chilean school meals auction. An important methodological contribution in our work is to introduce a reduced dimensional markup model in which bidders are assumed to determine their markups based on a reduced set of package characteristics. Our modeling approach is essential to achieve computational tractability for estimation in a large-scale CA.

We find that cost synergies in the Chilean school meals 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 also find that the effect of some of the side constraints currently used in the CA, which limit the market share each bidder can get in order to promote suppliers' diversification, results in only a small efficiency loss. We also compared the performance of the VCG mechanism to the first price CA used on this application. Contrary to results obtained in previous theoretical work where VCG has been criticized for achieving high procurement costs, we find that the total VCG payment is reasonable and quite close to the first-price CA payment.

Overall our results provide useful insights for the design of the Chilean auction. More broadly, our results highlight that the simultaneous consideration of the firms' operational cost structure and their strategic behavior is key to the successful design of a CA. Moreover, our structural estimation framework is sufficiently general to be used in other applications of large-scale CAs. In this way, we hope that this research agenda enhances the understanding of the performance of CAs and thereby provide insights to improve their design.

[^12]
## References

Aksin-Karaesmen, Z., B. Ata, S. Emadi, C. Su. 2011. Structural estimation of callers' delay sensitivity in call centers. Working Paper.

Allon, G., A. Federgruen, M. Pierson. 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 Operations Management 13(4) 489 - 507.

Athey, S., P. A. Haile. 2006. Empirical models of auctions. Invited lecture for the Ninth World Congress of the Econometric Society.

Ausubel, L. M., P. Cramton, R. P. Mcafee, J. Mcmillan. 1997. Synergies in wireless telephony: Evidence from the broadband pcs auctions. Journal of Economics and Management Strategy 6(3) 497-527.

Ausubel, L. M., P. Milgrom. 2006. The lovely but lonely Vickrey auction. P. Cramton, Y. Shoham, R. Steinberg, eds., Combinatorial Auctions. MIT Press, 17-40.

Cantillon, E., M. Pesendorfer. 2006. Combination bidding in multi-unit auctions. Working Paper.
Chapman, J.T.E., D. McAdams, H.J. Paarsch. 2005. Multi-unit, sealed-bid, discriminatory-price auctions. Working Paper, University of Iowa.

Cramton, P., Y. Shoham, R. Steinberg. 2006. Combinatorial Auctions. MIT Press.

Day, R., P. Milgrom. 2008. Core-selecting package auctions. International Journal of Game Theory 36(3-4) 393-407.

Day, R., S. Raghavan. 2007. Fair payments for efficient allocations in public sector combinatorial auctions. Management Science 53(9) 1389-1406.

Epstein, R., L. Henríquez, J. Catalán, G. Y. Weintraub, C. Martínez. 2002. A combinational auction improves school meals in Chile. Interfaces 32(6) 1-14.

Fox, J. T., P. Bajari. 2011. Measuring the efficiency of an FCC spectrum auction. Working Paper.

Gandal, N. 1997. Sequential auctions of interdependent objects: Israeli cable television licenses. The Journal of Industrial Economics XLV(3) 227-244.

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

Hendricks, K., R. H. Porter. 2007. An empirical perspective on auctions. M. Armstrong, R. Porter, eds., Handbook of Industrial Organization, Volume 3. Elsevier.

Horn, R. A., C. R. Johnson. 1985. Matrix Analysis. Cambridge University Press.

Hortaçsu, A., D. McAdams. 2010. Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the turkish treasury auction market. Journal of Political Economy 188(5) 833-865.

Kastl, J. 2011. Discrete bids and empirical inference in divisible good auctions. Review of Economic Studies 78(3) 974-1014.

Li, J., N. Granados, S. Netessine. 2011. Are consumers strategic? structural estimation from the air-travel industry. Working Paper.

Milgrom, P. 2000. Putting auction theory to work: The simultaneous ascending auction. Journal of Political Economy 108(2) 245 - 272.

Milgrom, P. 2004. Putting Auction Theory to Work. Cambridge.

Moreton, P. S., P. T. Spiller. 1998. What's in the air: Interlicense synergies in the federal communications commission's broadband personal communication service spectrum auctions. The Journal of Law and Economics XLI 677-716.

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

Olivares, M., G. Y. Weintraub, R. Epstein, D. Yung. 2011. Combinatorial auctions for procurement: An empirical study of the chilean school meals auction. Forthcoming, Management Science.

Paarsch, H. J., H. Hong. 2006. An Introduction to the Structural Econometrics of Auction Data. MIT Press.

Reguant, M. 2011. The welfare effects of complementary bidding mechanisms: An application to electricity markets. Working Paper.

Wooldridge, J.M. 2002. Econometric analysis of cross section and panel data. The MIT press.

## Tables and Figures

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

Table 1 - Results from the first step regression (equation (8)) for 2003 auction. Robust standard errors are shown in parenthesis. Combination/Cluster volume is measured in million meals per year.

| Correlation Coefficients |  |  |  |  |  |  |
| ---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Region | 4 | 5 | 9 | 12 | 13 | Std. Dev. |
| 4 | $1.00(0.00)$ | $0.52(0.21)$ | $0.31(0.27)$ | $0.45(0.24)$ | $0.67(0.17)$ | $14.56(3.20)$ |
| 5 | - | $1.00(0.00)$ | $0.65(0.16)$ | $0.69(0.17)$ | $0.69(0.13)$ | $14.52(2.55)$ |
| 9 | - | - | $1.00(0.00)$ | $0.42(0.22)$ | $0.09(0.27)$ | $22.92(4.02)$ |
| 12 | - | - | - | $1.00(0.00)$ | $0.48(0.22)$ | $46.48(9.97)$ |
| 13 | - | - | - | - | $1.00(0.00)$ | $13.46(2.29)$ |

Table 2 - Results from the second step regression (equation (9)) for 2003 auction. Standard errors are shown in parenthesis. Standard deviations of regional effects are measured in Chilean Pesos.

| Firm | Prob | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ | 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 |

Table 3 - Results from the markup estimation for representative firms of different winning probability levels for 2003 auction. Prob refers to the probability that the firm wins any package. The rest are the average per-meal markups corresponding to each package size. The markups are shown in the percentage of the average bid price per meal (US\$ 0.88).

| Size | $\mathbf{2}$ | $\mathbf{3}$ | $\mathbf{4}$ | $\mathbf{5}$ | $\mathbf{6}$ | $\mathbf{7}$ | $\mathbf{8}$ | Average |
| ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: | ---: |
| Cost Synergy (CH \$) | 5.13 | 10.64 | 14.40 | 13.71 | 15.07 | 16.64 | 18.93 | 12.82 |
| \% of Average Bid Price | 1.28 | 2.66 | 3.60 | 3.43 | 3.77 | 4.16 | 4.73 | 3.21 |
| Number of Observations | 289 | 87 | 121 | 49 | 126 | 169 | 205 |  |

Table 4 - Average cost synergies computed directly from estimated costs of individual units and multi-unit packages for 2003 auction. Size refers to the number of units in the package. The cost synergy measures the average per-meal cost savings when the units are combined to form a package of given size.


Volume of Combination (million meals/year)

Figure 1 - Average scale discounts in per-meal bid prices placed during 1999-2005 (Olivares et al., 2011). Dashed lines indicate $95 \%$ confidence interval of the estimates. The discount is measured by the decrease in the per-meal bid price when individual units are combined into a multi-unit package in the corresponding volume. For example, increasing package size to 20 million meals (combining about 8 units) generates discounts of around $6 \%$ of average bid price. All the bid prices are normalized to 1999 values using consumer price index.


Figure 2 - Estimates of the discount curves (equation (8)) and the synergy curves (equation (10)) for each group of firms for 2003 auction. Scale discounts and synergies are drawn against combination size and density discounts and synergies are drawn against cluster size. The estimation is done with 5101 observations.


Volume of Combination (million meals/year)

Figure 3 - Estimates of the discount curves (equation (8)) and the synergy curves (equation (10)) without density terms for 2003 auction. $95 \%$ confidence intervals are shown in dashed lines. The number of observations is 5101 .

## Online Appendix

## A Winner Determination Problem Formulation

In this section, we will provide the details of the integer programing (IP) formulation of the winner determination problem (WDP). This IP is formulated and solved in the course of simulation described in Section 4.2 to estimate each firm's markup vector. We begin by introducing notation that is not defined in the main body of the paper and we then formulate the IP.
Index Sets. We let $R$ denote the set of geographical regions indexed by $r$ (recall that each geographical region contains several TUs). We let $\mathcal{A}_{f}$ be the set of packages on which firm $f$ places bids. They are to distinguish from $\mathcal{A}$ in case of missing bids (unobserved bids) by firm $f$. $\mathcal{A}_{r f} \subseteq \mathcal{A}_{f}$ represents the set of packages in $\mathcal{A}_{f}$ that contain at least one TU in region $r$. Finally, we let $|a|$ denote the number of TUs in package $a$, and we let $A_{f}$ and $A_{r f}$ denote the number of packages in the sets $\mathcal{A}_{f}$ and $\mathcal{A}_{r f}$, respectively.
Decision Variables. We let $x_{a f}$ be the firm-package allocation decision variable for package $a$ and firm $f$. This variable takes the value of 1 , if firm $f$ wins package $a$, and 0 otherwise. These variables determine the final allocation. The variable $y_{r f}$ is a regional allocation variable for region $r$ and firm $f$, taking the value of 1 if firm $f$ wins a package that contains at least one TU in region $r$, and 0 otherwise. They are used to count the number of firms serving in each geographical region for the local constraints. The decision variable $z_{f}$ relates to the winning status of firm $f$. It is equal to 1 if firm $f$ wins a package and 0 otherwise. They count the number of winning firms to be used in the global competition constraint.
Constraints and Their Parameters. As described in Section 3, we have five types of allocative constraints in the auction. We also have an additional constraint imposed in our structural model, namely, that each firm can win at most one package. We label those constraints as follows: (A) Cover all TUs ensures that all the TUs be contracted. (B) At most one package constraint imposes that firms can win at most one package. (C) Maximum number of TUs bounds the number of TUs that each firm can win. We let $M X U_{f}$ denote the maximum number of TUs that firm $f$ can win. (D) Global Market Share Constraints limits the total volume of standing contracts of each firm in terms of the number of meals served. We let $M X M_{f}$ denote the total number of meals that firm $f$ can win in the auction being considered. (E) Local constraints bound the minimum and maximum number of firms serving in each region. We use $M N F_{r}$ and $M X F_{r}$ to denote these bounds for region $r$. (F) Global competition constraint sets the minimum number of firms being contracted in the auction being considered. We let $M N F_{g}$ denote this minimum number.

Notice that constraints (C) and (D) are firm-wise limits, and for each firm any bids placed on packages that exceed the firm's limits can never win. Therefore, we eliminate such bids a priori from $\mathcal{A}_{f}$ for each firm $f \in F$. That is for any given firm $f$ and for all $a \in \mathcal{A}_{f}$, we have $|a| \leq M X U_{f}$ and $v_{a} \leq M X M_{f}$. Then, constraints (C) and (D) will be automatically satisfied as long as firms win at most one package imposed by (B). Hence, we omit (C) and (D) in our IP formulation. Recall that the objective is to minimize the total procurement cost.

Now we present the IP formulation of the WDP. The constraints that are not labeled impose the correct values for the auxiliary variables $y_{r f}$ and $z_{f}$, and the integrality constraints for all decision variables.

$$
\begin{aligned}
\text { minimize } & \sum_{f \in F} \sum_{a \in \mathcal{A}_{f}} b_{a f} x_{a f} \\
\text { subject to (A) } \quad & \sum_{f \in F} \sum_{a \in \mathcal{A}_{f}: i \in a} x_{a f} \geq 1, \quad \forall i \in U \\
\text { (B) } \quad & \sum_{a \in \mathcal{A}_{f}} x_{a f} \leq 1, \quad \forall f \in F \\
\text { (E) } \quad & M N F_{r} \leq \sum_{f \in F} y_{r f} \leq M X F_{r}, \quad \forall r \in R \\
& \frac{1}{A_{r f}} \sum_{a \in \mathcal{A}_{r f}} x_{a f} \leq y_{r f} \leq \sum_{a \in \mathcal{A}_{r f}} x_{a f}, \quad \forall r \in R, \forall f \in F \\
& \sum_{f \in F} z_{f} \geq M N F_{g}, \\
& \frac{1}{A_{f}} \sum_{a \in \mathcal{A}_{f}} x_{a f} \leq z_{f} \leq \sum_{a \in \mathcal{A}_{f}} x_{a f}, \quad \forall f \in F \\
& x_{s f}, y_{r f}, z_{f} \in\{0,1\} .
\end{aligned}
$$

## B Proofs

## B. 1 Notation

We begin by defining notation that is frequently used in this section. First we consider a focal bidder $f$, whose observed bid vector is $b$. All of the analysis is focused on this bidder, and as before we omit the index $f$ whenever it is clear from the context. Recall that from the perspective of this focal bidder, competitors' bid prices are random. All such random quantities are defined over a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Note that $\mathbf{P}$ measures the probability of each of the events characterized by the final allocation of units to bidders in the CA. Hence, it defines the vector of winning probabilities $G(\cdot)$. In addition, we define $\Omega^{*} \subseteq \Omega$ to be the sample space where ties never happen in the winner determination problem. By Assumption 4, the distribution of competitors' bids is absolutely continuous, and hence, we can find such a sample space so that $\mathbf{P}\left(\Omega^{*}\right)=1$. In words, this means that the winner determination problem has a unique solution for any realization $\omega \in \Omega^{*}$. Therefore, in our analysis we do not consider any issues related to tie-breaking in the final allocation.

We let $b^{\prime}$ be the vector of competitors' bid prices. That is, given a realization of $\omega \in \Omega^{*}, b^{\prime}(\omega)=$ $\left\{b_{f^{\prime}}^{\prime}(\omega)\right\}_{f^{\prime} \neq f}$, where $b_{f^{\prime}}^{\prime}(\omega)$ is a vector of bids for competing firm $f^{\prime}$. Furthermore, we let $x=\left\{x_{a f}\right\}_{a \in \mathcal{A}, f \in F}$ be a $A \times|F|$ dimensional vector such that $x_{a f}$ takes 1 if bidder $f$ wins package $a$ and 0 otherwise. A vector $x$ uniquely determines an allocation outcome. We denote by $X$, the set of all feasible allocation outcomes that satisfy all the allocative constraints in the CA including the one that each bidder can win at most one package (see Assumption 1). All the proofs in this section are valid under any other allocative constraints in the CA as long as they do not depend on bid prices (so the constraints described in Section 3.2 are all valid). In addition, we let $X_{a} \subset X$ be the set of allocations such that bidder $f$ wins package $a$. We adopt
the null package indexed by 0 , and accordingly, we use $X_{0}$ to denote the set of allocations in which bidder $f$ wins no package, and let $G_{0}(b)$ be the probability that bidder $f$ wins no package given her bid vector $b$. Note that because bidders can win at most one package, $X_{a}$ and $X_{s}$ are disjoint for any packages $a \neq s$ in $\mathcal{A}_{0}=\mathcal{A} \cup\{0\}$, and we have $\bigcup_{a \in \mathcal{A}_{0}} X_{a}=X$.

Without loss of generality, we assume $x$ is ordered in a way such that the vector of bidder $f$ 's allocation decisions, denoted by $x_{f}$, is followed by the vector of competitors' allocation decisions, denoted by $x^{\prime}$, so that $x=\left(x_{f}, x^{\prime}\right)$. Additionally, we define a cost function: $p_{a}(\omega):=\min _{x \in X_{a}}\left(b, b^{\prime}(\omega)\right)^{T} x$, for each $a \in \mathcal{A}_{0}$. This is the minimum total procurement cost out of all the allocations such that bidder $f$ wins package $a$ given a realization $\omega \in \Omega^{*}$. It is important to note that because of the constraint that each bidder can win at most one package, $p_{a}(\omega)$ only depends on the value of $b_{a}$ among bidder $f$ 's bids for all $a \in \mathcal{A}$.

Finally, for notational simplicity, we use $G_{a, s}(b)$ to denote the partial derivative of the winning probability $G_{a}(b)$ with respect to the bid price $b_{s}$. In addition, when dealing with a characteristic-based markup model, we let $\mathcal{A}_{i} \subseteq \mathcal{A}$ to denote the set of packages associated with the $i^{\text {th }}$ markup variable, and also let $G_{a, \theta_{i}}(b)$ to denote the partial derivative of the winning probability $G_{a}(b)$ with respect to the markup variable $\theta_{i}$.

## B. 2 Proofs

We will use some lemmas for the proofs of the main results. The following lemma is useful for the proof of Proposition 1. The result follows by applying the fundamental theorem of calculus and its proof is omitted.

Lemma B.1. Define a function $F: \Re^{n} \mapsto \Re$ such that:

$$
F(y)=\int_{D(y)} f(x) d x
$$

where $f: \Re^{m} \mapsto \Re$ is continuous and integrable in $\Re^{m}$. Assume that the domain of integration $D(y)$ is a polyhedron formed by a given matrix $A \in \Re^{k \times m}$ and a vector function $b(y) \in \Re^{k}$ with $k \in \mathbb{N}$ such that $D(y):=\left\{x \in \Re^{m}: A x \leq b(y)\right\}$. If $b(y)$ is differentiable with respect to $y$, then $F$ is continuous and differentiable everywhere in $\Re^{n}$.

Proof of Proposition 1. To prove the differentiability of the winning probability vector $G(b)$ with respect to $b$, we begin by considering an arbitrary package $a \in \mathcal{A}$ and look at the winning probability that bidder $f$ wins package $a, G_{a}(b)$. Notice that bidder $f$ wins package $a$ if one of the allocations in $X_{a}$ achieves the minimum procurement cost among all possible allocations in $X$. We let $K:=\left|X_{a}\right|$, the number of distinct allocations in $X_{a}$, and index them by $k=1,2, \cdots, K$. Specifically, we look at the event that bidder $f$ wins package $a$ as a result of allocation $x_{k} \in X_{a}$. Accordingly, we let $G_{a}\left(b ; x_{k}\right)$ denote the probability that $x_{k} \in X_{a}$ becomes the final allocation (hence the minimizer of the total procurement cost). Because the probability of ties is zero, the winning probability $G_{a}(b)$ can be expressed as $G_{a}(b)=\sum_{k=1}^{K} G_{a}\left(b ; x_{k}\right)$. Therefore it suffices to show that $G_{a}\left(b ; x_{k}\right)$ is continuous and differentiable for any given allocation $x_{k}$.

Now given an arbitrary allocation $x_{k} \in X_{a}$, we show the differentiability of $G_{a}\left(b ; x_{k}\right)$ using Lemma B.1. By letting $f\left(b^{\prime}\right)$ denote the joint probability density function of competitors' bids $b^{\prime}$, the winning probability $G_{a}\left(b ; x_{k}\right)$ can be written as: $G_{a}\left(b ; x_{k}\right)=\int_{D_{k}(b)} f\left(b^{\prime}\right) d b^{\prime}$, where $D_{k}(b)$ is the set of $b^{\prime}$ 's for
which $x_{k}$ is the optimal allocation given $b$. Observe that $D_{k}(b)$ can be expressed by a set of inequalities as follows:

$$
x_{k}^{T}\left(b, b^{\prime}\right) \leq y^{T}\left(b, b^{\prime}\right), \quad \forall y \in X \quad(\Rightarrow) \quad\left(x_{k}^{\prime}-y^{\prime}\right)^{T} b^{\prime} \leq\left(y_{f}-x_{k f}\right)^{T} b, \quad \forall y \in X
$$

The inequalities ensure that, given the placed bids $\left(b, b^{\prime}\right)$, the total procurement cost incurred by allocation $x_{k}$ is cheaper than that of any other feasible allocations if we do not consider ties. Therefore, we get: $D_{k}(b)=\left\{b^{\prime} \in \Re^{A \times(|F|-1)}:\left(x_{k}^{\prime}-y^{\prime}\right)^{T} b^{\prime} \leq\left(y_{f}-x_{k f}\right)^{T} b, \forall y \in X\right\}$. If we let $J=|X|$ and index the feasible allocations by $j$, then $D_{k}(b)$ is a polyhedron in $\Re^{A \times(|F|-1)}$ defined by $M b^{\prime} \leq h(b)$, where the $j^{t h}$ row of $M$ is $\left(x_{k}^{\prime}-y_{j}^{\prime}\right)^{T}$ and the $j^{t h}$ element of vector $h(b)$ is $\left(y_{f}-x_{k f}\right)^{T} b$, for $j=1, \cdots, J$.

By Assumption 4, the density $H\left(\cdot \mid Z_{f^{\prime}}\right)$ is continuous everywhere and independent across bidders, and hence, the joint density $f\left(b^{\prime}\right)$ is continuous on $\Re^{A \times(F-1)}$. The integrability of $f\left(b^{\prime}\right)$ is readily obtained as it is a probability density function. Finally, the function $h(b)$ is a linear function of $b$, and hence differentiable with respect to $b$. Therefore, by Lemma B.1, $G_{a}\left(b ; x_{k}\right)$ is continuous and differentiable with respect to the bid vector $b$. Since the choice of package $a \in \mathcal{A}$ and allocation $x_{k} \in X_{a}$ was arbitrary, the proof is complete.

It is useful to examine some of the properties of the Jacobian matrixes $\mathcal{D}_{b} G(b)$ and $\mathcal{D}_{\theta} W^{T} G(b)$ for the proof of Proposition 2 and Theorem 1. The following lemma investigates those properties.

Lemma B.2. For any given bidder and her bid vector b, we have the following properties for the winning probability vector $G(b)$.

1. The Jacobian matrix $\mathcal{D}_{b} G(b)$ is symmetric.
2. For any package $a \in \mathcal{A}$, we have i) $G_{a, a}(b) \leq 0$; ii) $G_{s, a}(b) \geq 0$ for all $s \in \mathcal{A} \backslash\{a\}$; and iii) $\sum_{s \in \mathcal{A}} G_{s, a}(b) \leq 0$.
3. Consider a group markup model specified by a package-characteristic matrix $W$ whose elements are all non-negative. Let the markup vector $\theta$ and $D:=\mathcal{D}_{\theta} W^{T} G(b)$. Then $D_{i j} \geq 0$ for any $i \neq j$.

Proof. (Proof of Part 1). We fix two arbitrary but distinct packages $a$ and $s$, and we first show that $G_{s, a}(b) \leq G_{a, s}(b)$. We then establish the reversed inequality by exchanging the two packages and using a symmetric argument. The arbitrary choice of the two packages $a$ and $s$ then provides the completion of the proof.

Accordingly, take any two distinct packages $a, s \in \mathcal{A}$ and an arbitrary scalar $\epsilon>0$. We begin by defining the following events:

$$
\begin{aligned}
\Omega_{a} & :=\left\{\omega \in \Omega^{*}: p_{a}(\omega)=\min _{t \in \mathcal{A}_{0}} p_{t}(\omega)\right\} \\
\Omega_{a, s} & :=\left\{\omega \in \Omega_{a}: p_{s}(\omega)=\min _{t \in \mathcal{A}_{0} \backslash\{a\}} p_{t}(\omega)\right\} \\
\Omega_{a, s}^{\epsilon} & :=\left\{\omega \in \Omega_{a}: p_{s}(\omega)<p_{a}(\omega)+\epsilon\right\}
\end{aligned}
$$

By definition, $\Omega_{a}$ denotes the event where bidder $f$ wins package $a$, and $\Omega_{a, s} \subset \Omega_{a}$ denotes the events where the minimum allocation without bidder $f$ winning package $a$ is the one with her winning package $s$.

Also, $\Omega_{a, s}^{\epsilon} \subset \Omega_{a}$ is the event where the minimum allocation with bidder $f$ winning package $s$ is less than $\epsilon$ above from the optimal value, $p_{a}(\omega)$. Finally we let $\Omega_{s} \subset \Omega^{*}$ to be the event where bidder $f$ wins package $s$. Note that $\Omega_{a}$ and $\Omega_{s}$ are disjoint.

We use the following random variables: $Y_{s}^{a \pm \epsilon}(\omega):=\mathbf{1}\left[p_{s}(\omega)=\min \left(\min _{t \in \mathcal{A}_{0} \backslash\{a\}} p_{t}(\omega), p_{a}(\omega) \pm \epsilon\right)\right]$. The random variables $Y_{s}^{a \pm \epsilon}(\omega)$ indicate bidder $f$ 's winning of package $s$ when her bid $b_{a}$ changes by $+\epsilon$ and $-\epsilon$, respectively. Similarly, we define $Y_{s}^{0}(\omega):=\mathbf{1}\left[p_{s}(\omega)=\min _{t \in \mathcal{A}_{0}} p_{t}(\omega)\right]$, that is, the indicator that the bidder wins package $s$ given her bid price $b$ at the realization of $\omega$. Now we divide the event set $\Omega^{*}$ into the following four disjoint subsets and examine the values of the random variables $Y_{s}^{a+\epsilon}(\omega)$ and $Y_{s}^{0}(\omega)$.

1. $\forall \omega \in \Omega^{*} \backslash\left(\Omega_{a} \cup \Omega_{s}\right)$ : Bidder $f$ is winning neither $a$ nor $s$, so $Y_{s}^{0}(\omega)=0$. Moreover, increasing her bid $b_{a}$ by $\epsilon$ will not let her win $s$, hence, $Y_{s}^{a+\epsilon}(\omega)=0$.
2. $\forall \omega \in \Omega_{s}$ : Bidder $f$ is winning package $s$ and increasing her bid on non-winning package $a$ will not change her winning $s$. Thus, $Y_{s}^{0}(\omega)=Y_{s}^{a+\epsilon}(\omega)=1$.
3. $\forall \omega \in \Omega_{a, s} \cap \Omega_{a, s}^{\epsilon}$ : Bidder $f$ is winning package $a$, so $Y_{s}^{0}(\omega)=0$. Since $\omega \in \Omega_{a, s}^{\epsilon}$, after increasing $b_{a}$ by $\epsilon$, the value of the current optimal allocation $p_{a}(\omega)+\epsilon$ becomes larger than $p_{s}(\omega)$. But then, $\omega \in$ $\Omega_{a, s}$ implies $p_{s}(\omega)$ becomes the lowest cost allocation after such a perturbation. Hence, $Y_{s}^{a+\epsilon}(\omega)=1$.
4. $\forall \omega \in \Omega_{a} \backslash\left(\Omega_{a, s} \cap \Omega_{a, s}^{\epsilon}\right)$ : Bidder $f$ is winning package $a$, so $Y_{s}^{0}(\omega)=0$. If $\omega \notin \Omega_{a, s}$, after increasing $b_{a}$ by $\epsilon, p_{s}(\omega)$ is not the lowest cost allocation. If $\omega \notin \Omega_{a, s}^{\epsilon}, p_{s}(\omega)$ is still larger than the value of the current allocation, $p_{a}(\omega)+\epsilon$, even after the perturbation. Hence, $Y_{s}^{a+\epsilon}(\omega)=0$.

In words, $\left(\Omega_{a, s} \cap \Omega_{a, s}^{\epsilon}\right)$ is the only event in which bidder $f$ 's winning status of package $s$ changes by an $\epsilon$ increase in her bid $b_{a}$. Therefore, we obtain:

$$
\begin{equation*}
\frac{G_{s}\left(b+\epsilon e_{a}\right)-G_{s}(b)}{\epsilon}=\frac{1}{\epsilon} \mathbf{E}\left[Y_{s}^{a+\epsilon}-Y_{s}^{0}\right]=\frac{1}{\epsilon} \mathbf{P}\left(\Omega_{a, s} \cap \Omega_{a, s}^{\epsilon}\right), \tag{B.1}
\end{equation*}
$$

where $e_{a}$ is the $a^{\text {th }}$ canonical vector whose $a^{\text {th }}$ component is the only non-zero element and is equal to one.
Now we look at the effect of decreasing $b_{s}$ by $\epsilon$ to the winning of package $a$. Similarly, we divide the event set $\Omega^{*}$ into the following three disjoint subsets and examine the values of random variables $Y_{a}^{s-\epsilon}(\omega)$ and $Y_{a}^{0}(\omega)$.

1. $\forall \omega \in \Omega^{*} \backslash\left(\Omega_{a}\right)$ : Since bidder $f$ is not winning package $a, Y_{a}^{0}(\omega)=0$. Moreover, decreasing her bid $b_{s}$ by $\epsilon$ will never let her win package $a$, hence, $Y_{a}^{s-\epsilon}(\omega)=0$.
2. $\forall \omega \in \Omega_{a, s}^{\epsilon}$ : Bidder $f$ is winning package $a$, so $Y_{a}^{0}(\omega)=1$. Since $\omega \in \Omega_{a, s}^{\epsilon}$, after decreasing $b_{s}$ by $\epsilon, p_{s}(\omega)-\epsilon$ has a lower cost than the current optimal value, $p_{a}(\omega)$, so bidder $f$ will win package $s$ instead of $a$. Hence, $Y_{a}^{s-\epsilon}(\omega)=0$.
3. $\forall \omega \in\left(\Omega_{a} \backslash \Omega_{a, s}^{\epsilon}\right)$ : Bidder $f$ is winning package $a$, so $Y_{a}^{0}(\omega)=1$. Since $\omega \notin \Omega_{a, s}^{\epsilon}$, decreasing $b_{s}$ by $\epsilon$ cannot make the value $p_{s}(\omega)-\epsilon$ better than the current optimal value, $p_{a}(\omega)$. Hence, the previous optimal allocation will remain optimal and $Y_{a}^{s-\epsilon}(\omega)=1$.

This time, $\Omega_{a, s}^{\epsilon}$ is the only case that bidder $f$ 's winning status of package $a$ is affected by an $\epsilon$ decrease in her bid $b_{s}$. Therefore:

$$
\begin{equation*}
\frac{G_{a}(b)-G_{a}\left(b-\epsilon e_{s}\right)}{\epsilon}=\frac{1}{\epsilon} \mathbf{E}\left[Y_{a}^{0}-Y_{a}^{s-\epsilon}\right]=\frac{1}{\epsilon} \mathbf{P}\left(\Omega_{a, s}^{\epsilon}\right) . \tag{B.2}
\end{equation*}
$$

Since $\left(\Omega_{a, s} \cap \Omega_{a, s}^{\epsilon}\right) \subseteq \Omega_{a, s}^{\epsilon}$, from (B.1) and (B.2) we get the following inequality:

$$
\frac{G_{s}\left(b+\epsilon e_{a}\right)-G_{s}(b)}{\epsilon}=\frac{1}{\epsilon} \mathbf{P}\left(\Omega_{a, s} \cap \Omega_{a, s}^{\epsilon}\right) \leq \frac{1}{\epsilon} \mathbf{P}\left(\Omega_{a, s}^{\epsilon}\right)=\frac{G_{a}(b)-G_{a}\left(b-\epsilon e_{s}\right)}{\epsilon} .
$$

Recall that $\epsilon$ is an arbitrary positive scalar and Proposition 1 ensures the differentiability of $G(b)$ with respect to $b$. Thus, by letting $\epsilon$ vanish, we get $G_{s, a} \leq G_{a, s}$.

In the previous argument, the only condition for the packages $a$ and $s$ is that they are distinct. Hence, a symmetric argument also holds true and we get $G_{s, a} \geq G_{a, s}$, and therefore we get $G_{s, a}=G_{a, s}$. The arbitrary choice of $a$ and $s$ then let us conclude $G_{a, s}=G_{s, a}$, for any two distinct packages $a, s \in \mathcal{A}$. This completes the proof of part 1 .
(Proof of Part 2). To show $G_{a, a}(b) \leq 0$, fix a realization of $\omega \in \Omega^{*}$ and consider a perturbation of increasing bidder $f$ 's bid price $b_{a}$ by $\epsilon>0$. If she currently wins package $a$, she may or may not win package $a$ after the perturbation. However, if she currently does not win package $a$, i.e., $p_{a}(\omega)$ is not the lowest cost allocation, she cannot win package $a$ after the perturbation since $p_{a}(\omega)+\epsilon$ remains being larger than the current optimal value. Since these are true for any $\omega \in \Omega^{*}$, increasing bid price $b_{a}$ will never increase her chances of winning package $a$. Hence we get $G_{a}\left(b+\epsilon e_{a}\right) \leq G_{a}(b)$, for all $\epsilon>0$. Then the differentiability of $G(b)$, shown in Proposition 1, implies $G_{a, a}(b) \leq 0$.

Similarly, for the proof of $G_{s, a}(b) \geq 0$ for any $s \in \mathcal{A} \backslash\{a\}$, consider a perturbation of decreasing $b_{a}$ by an arbitrary $\epsilon>0$. Given a realization of $\omega \in \Omega^{*}$, if she currently wins package $s$ (possibly the null package), she can either win package $a$ instead of $s$ or still win package $s$ after the perturbation. However, if she currently wins package $a$, she will win package $a$ for sure after the perturbation. Therefore, decreasing her bid $b_{a}$ only possibly decrease her chances of winning package $s$, and we get $G_{s}(b) \geq G_{s}\left(b-\epsilon e_{a}\right)$, for all $\epsilon>0$. Again by the differentiability of $G(b)$, we get $G_{s, a}(b) \geq 0$.

Finally, since $\sum_{s \in \mathcal{A}} G_{s}(b)=1-G_{0}(b)$, so we get $\sum_{s \in \mathcal{A}} G_{s, a}(b)=-G_{0, a}(b) \leq 0$, where the last inequality follows because $G_{0, a}(b) \geq 0$ by a similar argument than above. This completes the proof of part 2 .
(Proof of Part 3). Note that by Assumption 5, $b=W \theta+c$ and by the chain rule, we have $D:=$ $\mathcal{D}_{\theta} W^{T} G(b)=W^{T} \mathcal{D}_{b} G(b) W$. Then for any $i \neq j$ we get:

$$
D_{i j}=\sum_{a, s \in \mathcal{A}} W_{a i} W_{s j} G_{a, s}(b)=\sum_{a \in \mathcal{A}_{i}, s \in \mathcal{A}_{j}} W_{a i} W_{s j} G_{a, s}(b),
$$

where the second equality comes from the fact that $W_{a i}=0$ if $a \notin \mathcal{A}_{i}$ by its definition. In addition, recall that in the group markup model, $\mathcal{A}_{i}$ and $\mathcal{A}_{j}$ are disjoint if $i \neq j$. Therefore by part 2 of this lemma shown above, $G_{a, s}(b) \geq 0$ for all $a \in \mathcal{A}_{i}$ and $s \in \mathcal{A}_{j}$. The non-negativity of the elements in $W$ then ensures that $D_{i j} \geq 0$ for all $i \neq j$, which completes the proof of part 3 .

Proof of Proposition 2. In the full-dimensional markup model, we have $b_{a}=c_{a}+\theta_{a}$ for $a=1, \ldots, K$, and the first-order conditions, (3) yields:

$$
\begin{equation*}
\left[\mathcal{D}_{b} G(b)\right]^{T} \theta=-G(b), \text { where } \theta=\left[\theta_{1}, \ldots, \theta_{K}\right]^{T} \tag{B.3}
\end{equation*}
$$

Similarly, for the common markup specification, we have $b_{a}=c_{a}+\theta_{u}$ for all $a=1, \ldots, K$. Note that the package-characteristic matrix $W \in \Re^{K}$ is then $W=[1,1, \ldots, 1]^{T}$. By letting $\alpha:=\left[\alpha_{1}, \ldots, \alpha_{K}\right]^{T}$ where $\alpha_{a}:=G_{a, \theta_{u}}(b)$, we have $\mathcal{D}_{\theta_{u}} W^{T} G(b)=W^{T} \mathcal{D}_{\theta_{u}} G(b)=W^{T} \alpha$. Then the first-order condition of this characteristic-based markup model, (5) now becomes:

$$
\begin{equation*}
\left[\mathcal{D}_{\theta_{u}} W^{T} G(b)\right]^{T} \theta_{u}=-W^{T} G(b) \quad(\Rightarrow) \quad \alpha^{T} W \theta_{u}=-W^{T} G(b) \tag{B.4}
\end{equation*}
$$

Observe that by definition, $\frac{\partial b_{s}}{\partial \theta_{u}}=1$ for all $s=1,2, \ldots, K$. Therefore by the chain rule, we get:

$$
\alpha_{a}=G_{a, \theta_{u}}(b)=\sum_{s=1}^{K} G_{a, s}(b) \quad(\Rightarrow) \quad W^{T}\left[\mathcal{D}_{b} G(b)\right]^{T}=\alpha^{T}
$$

Using this, left-multiplying by $W^{T}$ on both sides of equation (B.3) and then equating the right-hand sides of equations (B.3) and (B.4) yields:

$$
\sum_{a=1}^{K} \alpha_{a} \theta_{a}=\left(\sum_{a=1}^{K} \alpha_{a}\right) \theta_{u} \quad(\Rightarrow) \quad \theta_{u}=\frac{1}{\sum_{a=1}^{K} \alpha_{a}} \sum_{a=1}^{K} \alpha_{a} \theta_{a}
$$

Finally, by symmetry of the Jacobian matrix $\mathcal{D}_{b} G(b)$, shown in part 1 of Lemma B.2, we have $\alpha_{a}=$ $\sum_{s=1}^{K} G_{a, s}=\sum_{s=1}^{K} G_{s, a}$. Then part 2 of the same lemma implies $\alpha_{a} \leq 0$ for all $a=1, \ldots, K$. We end the proof by showing that at least one $\alpha_{a}<0$. Assume for the purpose of contradiction that $\alpha_{a}$ 's are all zero. This implies that the sum of all the column vectors in the Jacobian matrix $\mathcal{D}_{b} G(b)$ is a zero vector and therefore they are not linearly independent. However, since all the bids have strictly positive winning probabilities, the Jacobian matrix $\mathcal{D}_{b} G(b)$ is invertible as will be shown in Theorem 1, hence a contradiction. Therefore, we conclude that at least one $\alpha_{a}$ is strictly negative, and this completes the proof.

Proof of Proposition 3. The first-order conditions of the full-dimensional model, (3) gives:

$$
\left[\begin{array}{ccc}
G_{1,1}(b) & G_{2,1}(b) & G_{12,1}(b)  \tag{B.5}\\
G_{1,2}(b) & G_{2,2}(b) & G_{12,2}(b) \\
G_{1,12}(b) & G_{2,12}(b) & G_{12,12}(b)
\end{array}\right]\left[\begin{array}{c}
\theta_{1} \\
\theta_{2} \\
\theta_{12}
\end{array}\right]=-\left[\begin{array}{c}
G_{1}(b) \\
G_{2}(b) \\
G_{12}(b)
\end{array}\right]
$$

Now consider the case where we use common markup $\theta_{u}$ for single unit bids and markup $\theta_{v}$ for the package of units 1 and 2 , so that the package-characteristic matrix $W$ is formed as follows:

$$
W=\left[\begin{array}{ll}
1 & 0 \\
1 & 0 \\
0 & 1
\end{array}\right] \quad \rightarrow \quad \begin{array}{ll}
\rightarrow & \text { Unit 1: apply unit markup } \theta_{u} \\
\rightarrow & \text { Unit 2: apply unit markup } \theta_{u} \\
\text { Package 12: apply package markup } \theta_{v}
\end{array}
$$

Note that by the chain rule, $G_{a, \theta_{u}}(b)=G_{a, 1}(b)+G_{a, 2}(b)$, for $a=1,2,12$. Using this, the first-order condition of the characteristic-based model, (5) yields:

$$
\left[\begin{array}{cc}
G_{1,1}(b)+G_{1,2}(b)+G_{2,1}(b)+G_{2,2}(b) & G_{12,1}(b)+G_{12,2}(b)  \tag{B.6}\\
G_{1,12}(b)+G_{2,12}(b) & G_{12,12}(b)
\end{array}\right]\left[\begin{array}{c}
\theta_{u} \\
\theta_{v}
\end{array}\right]=-\left[\begin{array}{c}
G_{1}(b)+G_{2}(b) \\
G_{12}(b)
\end{array}\right]
$$

By left-multiplying by $W^{T}$ on both sides of (B.5) and then equating the right-hand sides of equations (B.5) and (B.6), we get the desired relationship of the two markup vectors. Note that by Theorem 1, the Jacobian matrix in (B.6) is invertible and therefore det $\neq 0$. This completes the proof.

Proof of Lemma 1. Fix a package $a \in \mathcal{A}$. Note that by the chain rule and Assumption 5, we have $G_{a, \theta_{i}}(b)=\sum_{s \in \mathcal{A}} \frac{\partial b_{s}}{\partial \theta_{i}} G_{a, s}(b)=\sum_{s \in \mathcal{A}} W_{s i} G_{a, s}(b)$. Therefore, it suffices to show that $G_{a, s}(b)=0$ for all $s \in \mathcal{A}$.

First, we let $\underline{p}(\omega):=\min _{t \in \mathcal{A}_{0}} p_{t}(\omega)$, the minimum procurement cost given $\omega \in \Omega^{*}$. Note that $G_{a}(b)=$ 0 implies $p_{a}(\omega)>p(\omega)$ in a set of $\Omega_{a} \subseteq \Omega^{*}$, such that $\mathbf{P}\left(\Omega_{a}\right)=1$. Also, we let $e_{a} \in \Re^{A}$ be the $a^{\text {th }}$ canonical vector, whose $a^{\text {th }}$ component is equal to one while all others are equal to zero.

We now show that $G_{a, s}(b)=0$ for all $s \in \mathcal{A} \backslash\{a\}$. First, take any package $s \neq a$ and consider a perturbation of decreasing $b_{s}$ by $\epsilon>0$. Recall that bidder $f$ can win at most one package and therefore $p_{a}(\omega)$ does not depend on the value of $b_{s}$. Therefore, decreasing $b_{s}$ will not change the value of $p_{a}(\omega)$. However, depending on whether $b_{s}$ is part of the current optimal allocation or not, the value of the current optimal allocation may decrease by $\epsilon$ or stay the same $(\underline{p}(\omega)$ ) after the perturbation. Thus, after such a perturbation the value of the current allocation will still have a lower cost than $p_{a}(\omega)$. This implies that bidder $f$ remains not winning package $a$ for all $\omega \in \Omega_{a}$. Hence, we obtain $G_{a}(b)-G_{a}\left(b-\epsilon e_{s}\right)=0$ for all $\epsilon>0$. Then the differentiability of $G(b)$ established in Proposition 1 implies $G_{a, s}(b)=0$.

Similarly, to show that $G_{a, a}(b)=0$, consider a perturbation of increasing $b_{a}$ by $\epsilon>0$. Then again for all $\omega \in \Omega_{a}$, after such a perturbation, $p_{a}(\omega)$ only increases to be $p_{a}(\omega)+\epsilon$, and remains being larger than the optimal value $\underline{p}(\omega)$. Hence bidder $f$ can never win package $a$ after the perturbation, which implies $G_{a}\left(b+\epsilon e_{a}\right)-G_{a}(b)=0$ for all $\epsilon>0$. Again by Proposition 1, we obtain $G_{a, a}(b)=0$.

By combining these results, we finally get $G_{a, \theta_{i}}(b)=\sum_{s \in \mathcal{A}} W_{s i} G_{a, s}(b)=0$, which completes the proof.

The following Lemma provides invertibility conditions of a matrix, which is used to prove Theorem 1.
Lemma B. 3 (Theorem 6.1.10 in Horn and Johnson (1985)). A matrix $D \in \Re^{n \times n}$ is said to be strictly diagonally dominant, if it satisfies:

$$
\left|D_{i i}\right|>\sum_{j \neq i}\left|D_{i j}\right|, \quad \forall i=1,2, \ldots, n .
$$

If $D$ is strictly diagonally dominant, then $D$ is invertible.
Proof of Theorem 1: (Necessity). We first show that if the Jacobian matrix $\mathcal{D}_{\theta} W^{T} G(b)$ is invertible it must be that every markup variable has at least one relevant bid associated with it. For this, assume there exists a markup variable, say $\theta_{i}$, whose associated bids are all irrelevant. Now note that $\left[\mathcal{D}_{\theta} W^{T} G(b)\right]_{i j}=$
$\sum_{a \in \mathcal{A}_{i}} W_{a i} G_{a, \theta_{j}}(b)$. But then Lemma 1 implies that $G_{a, \theta_{j}}(b)=0, \forall a \in \mathcal{A}_{i}$, leading to $\left[\mathcal{D}_{\theta} W^{T} G(b)\right]_{i j}=$ 0 . Since this is true for any $j=1,2, \ldots, d$, the $i^{t h}$ row of Jacobian matrix $\mathcal{D}_{\theta} W^{T} G(b)$ will be a zero vector. Having a row of zeros implies that the matrix is not invertible. This completes the proof of necessity.

Proof of Theorem 1: (Sufficiency). We now show that if every markup variable has at least one relevant bid associated with it and the additional conditions in the statement of the theorem hold, then the Jacobian matrix $\mathcal{D}_{\theta} W^{T} G(b)$ evaluated at the observed bid vector $b$ is invertible, and therefore the markup vector $\theta$ is uniquely determined by equation (6). For notational simplicity, we let $D:=\mathcal{D}_{\theta} W^{T} G(b)$.

First, recall that in a group markup specification, for any package $a$, there is only one markup variable that is associated with it, say markup variable $\theta_{i}$. Then the profit that bidder $f$ makes from winning package $a$ is $W_{a i} \theta_{i}$. By assumption, $W_{a i} \theta_{i} \geq 0$ and $W_{a i} \geq 0$, for all packages $a$. Therefore, $\theta_{i} \geq 0$, for all $i$. We now proceed to show that $\theta_{i}$ is indeed strictly positive for all $i=1,2, \ldots, d$. By assumption, $\theta$ satisfies equation (5): $D^{T} \theta=-W^{T} G(b)$. For the purpose of contradiction, we fix $i$ and assume that $\theta_{i}$ is zero. We examine the $i^{\text {th }}$ equation in (5):

$$
\begin{equation*}
D_{i i} \theta_{i}+\sum_{j \neq i} D_{j i} \theta_{j}=-W_{i}^{T} G(b) . \tag{B.7}
\end{equation*}
$$

The first term on the left-hand side is zero by assumption. The second term is non-negative since we know that (i) $\theta_{j} \geq 0, \forall j$; and (ii) $D_{j i} \geq 0$ by part 3 of Lemma B.2. However, the right-hand side is strictly negative because there is at least one relevant bid, say $b_{a}$, that is associated with markup variable $\theta_{i}$, so that $W_{i}^{T} G(b) \geq W_{a i} G_{a}(b)>0$. Therefore it is impossible for $\theta$ to satisfy equation (5), which contradicts our assumption. Hence, $\theta_{i}>0$, for all $i$.

Now, we construct a diagonal matrix $\Theta$ so that $\Theta_{i i}=\theta_{i}$ for all $i=1,2, \ldots, d$. Because $\theta_{i}>0, \forall i$, it is clear that $\Theta$ is invertible. We now show that equation (5) implies that the matrix $D^{T} \Theta$ is strictly diagonally dominant, and therefore invertible by Lemma B.3. To see this, take any $i \in\{1,2, \ldots, d\}$, and consider the $i^{\text {th }}$ equation in (5) (see (B.7)), for which we know that its right-hand side is strictly negative. Therefore, using $\left[D^{T} \Theta\right]_{i j}=D_{j i} \Theta_{j j}=D_{j i} \theta_{j}$, we reach the following inequality:

$$
\left[D^{T} \Theta\right]_{i i}+\sum_{j \neq i}\left[D^{T} \Theta\right]_{i j}=-W_{i}^{T} G(b)<0 \quad(\Rightarrow) \quad \sum_{j \neq i}\left[D^{T} \Theta\right]_{i j}<-\left[D^{T} \Theta\right]_{i i}
$$

Recall that when $i \neq j$, we have $\left[D^{T} \Theta\right]_{i j}=D_{j i} \theta_{j} \geq 0$, and this implies $\sum_{j \neq i}\left|\left[D^{T} \Theta\right]_{i j}\right|<\left|\left[D^{T} \Theta\right]_{i i}\right|$. Since this is true for any $i=1,2, \ldots, d$, we conclude that $D^{T} \Theta$ is strictly diagonally dominant and hence invertible by Lemma B.3. Since $\Theta$ is also invertible, the invertibility of $D$ follows with $D^{-1}=\left(\Theta^{T} D\right)^{-1} \Theta^{T}$. The proof for sufficiency is now complete.

## C VCG Payment Rule and a Core Outcome.

## C. 1 VCG Payment Rule

First, we describe the payment rules of the VCG mechanism, which we then use to calculate total payments under VCG. Let $V(F)$ denote the value of the minimum-cost allocation that satisfies all constraints based on the reported bids of all firms in set $F$. Because VCG is truthful, these bids correspond to actual costs. In
addition, let $F^{*} \subseteq F$ be the set of firms who are awarded contracts in the VCG allocation and let $b_{a(f), f}$ be the bid price reported by firm $f \in F^{*}$ for her winning package $a(f)$ (in this notation $b_{a(f), f}$ represents the total value for the entire package, not the per-meal value). The VCG payment to winner $f \in F^{*}$, denoted by $P_{f}$, is computed as follows:

$$
P_{f}=V\left(F_{-f}\right)-\sum_{f^{\prime} \in F_{-f}^{*}} b_{a\left(f^{\prime}\right), f^{\prime}}
$$

where $F_{-f}=F \backslash\{f\}$ and $F_{-f}^{*}=F^{*} \backslash\{f\}$. The first term is the total value of reported bids in the optimal allocation that considers all bids except those from winning firm $f$. The second term is the total value of reported bids in the current VCG allocation (that includes firm $f$ ), except for the reported value of firm $f$ 's winning package. Hence, the payment to a winner is essentially the cost of providing the units she wins in the lowest cost allocation without her. Loosing bidders do not receive payments. The total procurement cost for the auctioneer under VCG is then obtained by summing up all such individual payments to winning firms.

## C. 2 Finding a Core Outcome Close to VCG

Now we turn our attention to the concept of a core outcome in a CA. Specifically we are interested in checking whether the VCG outcome lies in the core or whether it is close to it. We start by providing some useful definitions. We closely follow Day and Raghavan (2007); Day and Milgrom (2008) also provide a useful description of this material. First, we call the final allocation and the payments to bidders in a CA an outcome. Given an outcome, $\Gamma$, we call the set of winning bidders a coalition, $C_{\Gamma}$. An outcome $\Gamma$ is said to be blocked if there exists an alternative outcome $\bar{\Gamma}$ that generates strictly lower total procurement cost to the auctioneer and for which every bidder in $C_{\bar{\Gamma}}$ weakly prefers $\bar{\Gamma}$ to $\Gamma$. An efficient outcome $\Gamma$ that is not blocked, is called a core outcome. Note that if an outcome is not in the core, there is a group of bidders that have incentives to deviate from it and offer a better deal to the auctioneer.

In addition, a core outcome $\Gamma$ is called bidder-Pareto optimal if there is no other core outcome weakly preferred by every bidder in $C_{\Gamma}$. Day and Raghavan (2007) and Day and Milgrom (2008) propose auctions that find efficient, core, bidder-Pareto optimal outcomes. An attractive property of efficient core-selecting auctions that are also bidder-Pareto optimal is that they minimize the incentives to unilaterally misreport true costs among all core-selecting auctions. In this sense, these auctions have outcomes that are closest to VCG among all core outcomes. We use the algorithm proposed by Day and Raghavan (2007) to find a core outcome that is closest to VCG. ${ }^{15}$

## D Counterfactual Results for 2005 Auction

In this section, we provide the counterfactual results for the 2005 auction. First, we find that the allocation is also highly efficient in 2005. Recall from Section 4.3 that the total annual supplying cost in the first-price CA is US\$ 51.53 million. The total annual supplying cost of the minimum-cost allocation is US\$ 51.49

[^13]million over the set of relevant bid packages and US $\$ 50.70$ million over the set of expanded package sets. This gives about $1.6 \%$ of efficiency loss in the allocation by the first-price CA.

Second, the VCG mechanism also achieves very close total procurement cost to that of the first-price CA. The total annual procurement cost under VCG is computed to be US $\$ 53.5$ million, which is only $0.23 \%$ more expensive than the total procurement cost of US\$ 53.4 million under the first-price CA. This time, the VCG payments are even closer to the core payments with respect to the truthful bids. The difference of the total procurement costs between these two points is less than $0.03 \%$ in 2005. Moreover, the individual payments are also closer; two-thirds of the nine winners receive exactly the same payments in the core point as in VCG and the rest three receive payments that are no more than $0.7 \%$ apart. Hence, in 2005, the VCG outcome is also essentially in the core.

Finally, in 2005, we have a bit larger but still small efficiency loss incurred by the allocative constraints. We consider the loss due to the market share constraints and global competition constraints. The efficiency loss in the constrained auction is $2.8 \%$ compared to the minimum-cost allocation without those constraints. In 2005 the impact of the global competition is higher; it imposes a minimum of 9 winners out of 16 bidders in 2005; in 2003 it also imposed the same minimum but out of 20 bidders.


[^0]:    *Acknowledgements: We thank Rafael Epstein for the valuable insights he provided to us. We are thankful for the helpful comments received from Lanier Benkard, Estelle Cantillon, Awi Federgruen, Jeremy Fox, Ali Hortaçsu, Jakub Kastl, Paul Milgrom, Martin Pesendorfer, Mar Reguant, Michael Riordan, Richard Steinberg, Assaf Zeevi, and seminar participants at Caltech SISL, Columbia Business School, Chicago Booth, Duke Fuqua, Informs Conference, London School of Economics, MSOM Conference, Revenue Management Conference, Stanford MS\&E, Stanford Structural IO Lunch, and Utah Winter Operations Conference. We thank Daniel Yung for exemplary research assistance. We also thank the Social Enterprise Program, CIBER, and the W. Edwards Deming Center at Columbia Business School for their financial support.
    ${ }^{\dagger}$ Columbia Business School, Email: skim14@gsb.columbia.edu,
    ${ }^{\ddagger}$ Columbia Business School, Email: molivares@columbia.edu
    ${ }^{\text {§ }}$ Columbia Business School, Email: gweintraub@columbia.edu

[^1]:    ${ }^{1}$ We note that to obtain the result that $\alpha_{a} \leq 0, \forall a$, it is important to assume that $\left\{\theta_{a}\right\}_{a \in\{1, \ldots, K\}}$ and $\theta_{u}$ are absolute markups, i.e., markup values for corresponding units or packages. Hence, elements in the package-characteristic matrix $W$ are either zero or one.

[^2]:    ${ }^{2}$ 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.

[^3]:    ${ }^{3}$ The following two conditions are necessary for identification for a general package characteristic specification: (i) every markup variable has at least one relevant bid associated with it; and (ii) the package-characteristic matrix $W^{R}$ has full column rank. We have not explored general sufficient conditions for identification beyond the group markup model.

[^4]:    ${ }^{4}$ Olivares et al. (2011) report that when $\tilde{\delta}_{i f}$ 's are treated as parameters being estimated, the explanatory power of the regression is remarkably high with an R -square equals to 0.98 .

[^5]:    ${ }^{5}$ In small experiments we also found that modeling these discount parameters as random variables does not affect the cost estimates by much.

[^6]:    ${ }^{6}$ Point-identification of the implicit unit prices $\tilde{\delta}_{i f}$ can be obtained when each firm submits many bids containing each unit $i$. Our empirical application meet these requirements. For example, in 2003, 20 firms participated for 33 units and on average submitted around 2100 bids each.

[^7]:    ${ }^{7}$ Recall that in Proposition 2 we assumed that $\left\{\theta_{a}\right\}_{a \in\{1, \ldots, K\}}$ and $\theta_{u}$ are absolute markups. In the per-meal markup specification, we obtain absolute markups by multiplying them by the total meal volume in the package through the $W$ matrix. In this case, the $\alpha_{a}$ 's are not necessarily all negative a priori. However, in our estimations they turned out to be negative for most of firms and auctions, leading to weighted average markups when aggregated.

[^8]:    ${ }^{8}$ We run 100,000 simulation runs for each firm. On average, it takes around 15 days to finish the simulation for the firms with relatively higher winning probabilities with about $10-18$ markup variables. For other firms, that have relatively lower winning probabilities, it took around 10 days with about 5 markup variables. The program is implemented in C with CPLEX V12.1 and ran on Columbia's research grid where each machine has eight 2.4 GHz CPUs. Note that to fully evaluate the local optimality of the markups, we need to estimate the Hessian matrices of the bidders' expected profit. However, this is computationally very intense, requiring an order of magnitude more computation time to what is required to estimate the markups.
    ${ }^{9}$ 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. Moreover, the number of firms and auctioned units were also larger, and we omit the results of this year as it requires an onerous amount of computational time. However, the estimation was more manageable for the 2005 auction as the number of units auctioned and the participating firms are smaller. We did not estimate years 1999-2002, because in those auctions bidders had less experience and history to rely on, and

[^9]:    were less sophisticated, so that our structural model assumptions may be harder to justify.
    ${ }^{10}$ There are two small firms whose estimated cost synergies are significantly different from the rest firms, and they are not accounted in the figure.

[^10]:    ${ }^{11}$ The results for the 2005 auction are similar and consistent with the 2003 auction. We provide the counterfactual results for the 2005 auction in Online Appendix D.

[^11]:    ${ }^{12}$ 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 ten firms in the efficient allocation; the majority of them -seven firms- are present in both cases. 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.

[^12]:    ${ }^{13}$ The final allocations in both cases look similar. Nine firms win in both cases and only one winner is replaced by another. Two firms win exactly the same packages, and six other firms have many of the winning units overlap in both cases or win units in the same region. The efficiency loss is mainly triggered by one large firm who won a package of two units with market share constraint and won a package of five units in the unconstrained case.
    ${ }^{14}$ Olivares et al. (2011) show that local competition, measured by the number of firms serving nearby units, has a significant effect in reducing prices in this application. This suggests that supplier diversification at a local level can lead to increased competition.

[^13]:    ${ }^{15}$ Note that a core-selecting auction may not be truthful, so in general it selects core outcomes with respect to the reported costs. In our analysis we restrict attention, however, to efficient core outcomes with respect to the truthful bids.

