Identification using Revealed Preferences in Linearly Separable Models*

Nikhil Agarwal, Pearl Z. Li and Paulo Somaini†

September 18, 2026

Abstract

Revealed preference arguments are commonly used when identifying models of both single-agent decisions and noncooperative games. We develop general identification results for a large class of models that have a linearly separable payoff structure. Our model allows for both discrete and continuous choices. It incorporates widely studied models such as discrete and hedonic choice models, auctions, school choice mechanisms, oligopoly pricing and trading games. We characterize the identified set and show that point identification can be achieved either in sufficiently rich choice environments or if a variable that shifts preferences is available. Our identification results also suggest an estimation approach. Finally, we implement this approach to estimate values in a combinatorial procurement auction for school lunches in Chile. Without assumptions on the distribution of costs, we find that an efficient auction would have reduced production costs by at least 3.8 percent of the payments to winning bidders, and roughly twice that under a parametric cost distribution.

*We are grateful to Gabriel Weintraub for sharing the data used in this project. We thank Brent Hickman, Eugenio Miravete, Francesca Molinari, Parag Pathak, Anna Russo, and Whitney Newey for helpful comments. Yucheng Shang and Miguel Borrero provided excellent research assistance. Agarwal acknowledges support from the Alfred P. Sloan Foundation (FG-2019-11484).

†Agarwal: Department of Economics, MIT and NBER, email: agarwaln@mit.edu. Li: Department of Economics, University of Illinois Urbana-Champaign, email: pearlzli@illinois.edu. Somaini: Stanford Graduate School of Business and NBER, email: soma@stanford.edu.

1 Introduction

Revealed preference arguments (Samuelson, 1938; Varian, 1982) are central to estimating the distribution of agents' payoffs. These arguments yield restrictions on preferences or payoffs based on an observed action (or choice). The typical argument assumes optimal behavior to derive implications of the requirement that any alternative action that an agent could have taken must yield a weakly lower payoff. The consequences of the chosen and alternative action – in terms of realized allocations – are then used to identify or bound the payoffs from various allocations.

Such arguments have been applied to study a seemingly disparate set of models. The most immediate application is to consumer choice (e.g., McFadden, 1973; Rosen, 1974).[1] However, arguments that are similar in spirit have been applied for other single-agent choice settings and a large class of noncooperative games. Single-agent examples include hedonic demand models (e.g., Rosen, 1974; Bajari and Benkard, 2005) and monopoly pricing (e.g., Luo et al., 2018). Examples of noncooperative games include games with incomplete information such as auctions (Guerre et al., 2000), school choice (Agarwal and Somaini, 2018), and trading or bargaining games (Larsen and Zhang, 2018); and games with full information such as oligopoly price setting where the objective is to identify marginal costs (Berry et al., 1995; Berry and Haile, 2014). A central question for these models is whether (the distribution of) agents' payoff types is identified from the available data.

Our starting point is the observation that several models in the literature share a common structure. An agent can take an action $a \in \mathcal{A}$. The consequence of the action is described by an expected outcome $x \in \mathcal{X} \subseteq \mathbb{R}^J$ and an expected transfer $t \in \mathbb{R}$. Payoffs take a linear form and the agent maximizes

$$v \cdot x_{\mathcal{A}}(a) - t_{\mathcal{A}}(a),$$

where $v \in \mathbb{R}^J$ is the agent's preference type and the functions $x_{\mathcal{A}}(\cdot)$ and $t_{\mathcal{A}}(\cdot)$ map actions to outcomes and transfers, respectively. We assume the analyst knows (or can identify) these functions and observes each agent's action. The goal is to identify $F_{V|Z}(v; z)$, the cumulative distribution function (CDF) of v given observables z, which is the model's primitive required for counterfactual analysis.

This paper studies the identification of models with this linearly separable structure. Our

[1] We abstract away from the endogeneity of prices which is the focus of an important literature on identifying market demand (Berry et al., 1995; Berry and Haile, 2014). We focus on cases where endogeneity can be perfectly controlled for by other means or where consumer-level price variation that is orthogonal to preferences is available (e.g. Tebaldi et al., 2019).

framework and results allow the induced outcome set $\mathcal{X}$ to be discrete, continuous, or a mixture of the two, and can incorporate both single-agent decision settings and a large class of noncooperative games with independent private information (see section 3). For example, the hedonic demand model of Rosen (1974) and Bajari and Benkard (2005) maps to the model above if a is the good chosen, $x_{\mathcal{A}}(a)$ denotes the characteristics of the good, v denotes the vector of consumer preferences for the characteristics and $t_{\mathcal{A}}(a)$ denotes the pricing function. In the class of incomplete information noncooperative games with independent and private types, $x_{\mathcal{A}}(a)$ and $t_{\mathcal{A}}(a)$ are expected outcomes and transfers, integrating over the strategies of other agents. For example, in the first-price auction studied in Guerre et al. (2000), a denotes the bid, $x_{\mathcal{A}}(a)$ denotes the probability of winning and $t_{\mathcal{A}}(a)$ denotes the expected payment.

The main theoretical results of the paper characterize the identified set of distributions $F_{V|Z}(v|z)$ and derive conditions under which it is point identified. Point identification can be achieved in two cases. The first is a sufficiently "rich" choice environment: the convexified transfer function t is differentiable at each observed outcome, so each observed outcome is compatible with a unique payoff type. In this case, our results also yield identification of the payoff type v_i for each agent i in the market. This result implies identification in the hedonic price model as well as the first-price independent private value auctions. It also implies identification of marginal costs in oligopoly models if the demand function is known.

In the complementary case, where each choice is optimal for a set of payoff types, we show conditions under which variation in an observable z can be used to "trace out" the distribution of v. In this case, we require that z acts as a payoff shifter as in Lewbel (2000), although we can allow for a more general nonlinear form by applying arguments similar to those in Allen and Rehbeck (2017). These results imply identification in discrete choice models, bargaining models with discrete offers (Larsen and Zhang, 2018), and school choice models with or without strategic manipulation (Agarwal and Somaini, 2018).

These results therefore unify the analysis of identification in a large class of models, which have thus far been obtained using arguments customized for each model. The sparse structure our results require is exactly what makes them portable, and we expect them to spur new specializations and further identification results. For instance, we allow for a combination of discrete and continuous choices, which have already been used in new contexts. Aspelund and Russo (2023), for instance, analyzes a scoring auction with multidimensional bids that include discrete and continuous components. In this paper, we use our framework to study high-dimensional combinatorial auctions.

Our characterization of the identified set also suggests an estimation approach that is portable across contexts. In a first step, an analyst can use the revealed preference restrictions that we derive to bound the payoff type of each agent in the dataset based on the agent's action. We show that these restrictions imply that the agent's payoff belongs to a convex set. The second step then estimates the distribution of payoffs using an estimator of choice and further restrictions (if any).

As an illustrative application of our approach, we empirically analyze the procurement for public school lunches in Chile that is based on a combinatorial auction. This auction was analyzed in Kim et al. (2014), henceforth KOW. Our study uses the same dataset. It is well-known that the extreme high dimensionality of the choice set in a combinatorial auction presents several technical challenges. We show how our reformulation of the problem suggests an alternative approach to the one taken in KOW.

The revealed-preference restrictions alone bound the aggregate margin on awarded bids below at 0.2 percent and imply that an efficient allocation of the observed bids would have reduced production costs by at least 15.8 pesos per meal, or 3.8 percent of the total payment to winning bidders. Adding a parametric restrictions on the cost distribution collapses the bounds to point estimates. With these additional distributional assumptions, we estimate that the efficient allocation saves 32.6 pesos per meal, or 8.6 percent of the observed allocation cost, a little more than twice the distribution-free lower bound. We remain agnostic about whether the additional parametric restrictions are plausible. Our primary point is that the approach can be used to analyze the model under various assumptions.

Related Literature

A large literature that is not easily summarized applies revealed preference arguments to show identification of various models. We point the reader to several surveys for identification results pertaining to these models; for example, see Berry and Haile (2016) for demand models; Athey and Haile (2007) for auction models; and Agarwal and Somaini (2020) for school choice.

Pakes (2010) uses revealed-preference inequalities to generate moment conditions that bound a finite-dimensional parameter, such as the parameters of a profit function or the coefficients from projecting expected profits on observables. Those inequalities remain valid when agents face uncertainty and make expectational errors. We ask instead what the inequalities identify about the full conditional distribution of payoff types, $F_{V|Z}$, of which those projection coefficients are finite-dimensional features. The two sets of results are non-nested. Like Ciliberto and Tamer (2009), our object of interest includes the distribution of payoff heterogeneity

rather than only selected moments of that distribution. Unlike their parametric approach, we characterize $F_{V|Z}$ nonparametrically and provide conditions for its point identification. We rely on payoff shifters as in Lewbel (2000) and generalize results on the identification of games with incomplete information presented by Lewbel and Tang (2015).

Methodologically, our identified set is an object from the theory of random sets. The optimal-choice correspondence is a random closed set and the selection kernels are its measurable selections, so the identified set is characterized by capacity-functional inequalities in the sense of Artstein (1983). This is the characterization used by Galichon and Henry (2011), dual to the Aumann-expectation characterization of Beresteanu et al. (2011). The source of incompleteness differs. In that literature, it typically stems from equilibrium selection in models with multiple equilibria, whereas here it stems from an incomplete model of tie-breaking among optimal actions. When the type distribution admits a density the model is complete, and the likelihood we use in our application is a coarse-data likelihood in the sense of Heitjan and Rubin (1991).

The illustrative application that we study is related to a small but growing literature on the analysis of combinatorial auctions. Papers that use bid data to estimate complementarities or substitutabilities between multiple objects that are auctioned off simultaneously include Cantillon and Pesendorfer (2007), Gentry et al. (2014) and Xiao and Yuan (2020).

Overview

Section 2 describes the notation of our model for single-agent problems and noncooperative games. Section 3 maps the leading examples to our notation. Section 4 presents the main results on identification and applies it to the examples. Section 5 presents the application to combinatorial auctions. Section 6 concludes. Proofs and lemmas not contained in the main text are in the appendix.

2 Model

2.1 Notation and Agent Decisions

Consider an agent indexed by i who picks an action a from a set $\mathcal{A}$.[2] Let $x_{\mathcal{A}} : \mathcal{A} \to \mathbb{R}^J$ and $t_{\mathcal{A}} : \mathcal{A} \to \mathbb{R}$ be the functions mapping actions to outcomes and transfers respectively, where $t_{\mathcal{A}}(\cdot)$ is bounded below. Each action results in an **outcome** described by $x \in \mathcal{X}$,

[2]The agent may randomize over actions in $\mathcal{A}$; the observed a is a realization of that mixture.

where $\mathcal{X} := \{x_{\mathcal{A}}(a) : a \in \mathcal{A}\}$, and an (expected) **payment** $t \in \mathbb{R}$. We do not restrict the cardinalities of either $\mathcal{X}$ or $\mathcal{A}$. By assuming that all agents face the same choice sets, our analysis effectively conditions on the specific choice set $(x_{\mathcal{A}}, t_{\mathcal{A}})$ faced by an agent.

Agent i's preference type is $v_i \in \mathbb{R}^J$. She chooses $a \in \mathcal{A}$ to maximize $v_i \cdot x_{\mathcal{A}}(a) - t_{\mathcal{A}}(a)$.

Optimality implies that an agent with preference type v picks action a only if for all $a' \in \mathcal{A}$,

$$v \cdot x_{\mathcal{A}}(a) - t_{\mathcal{A}}(a) \geq v \cdot x_{\mathcal{A}}(a') - t_{\mathcal{A}}(a').$$

Thus, if a and a' are such that $x_{\mathcal{A}}(a) = x_{\mathcal{A}}(a')$ and $t_{\mathcal{A}}(a) < t_{\mathcal{A}}(a')$, then no agent picks a'. Define $t_{\mathcal{X}}(x)$ to be cost of choosing the outcome $x \in \mathcal{X}$:

$$t_{\mathcal{X}}(x) = \inf\{t_{\mathcal{A}}(a),\ a \in \mathcal{A} : x_{\mathcal{A}}(a) = x\}.$$

Hence, $t_{\mathcal{X}} : \mathcal{X} \to \mathbb{R}$.[3] Observe that the action $a \in \mathcal{A}$ is dominated if $(x_{\mathcal{A}}(a), t_{\mathcal{A}}(a))$ is not in the graph of $t_{\mathcal{X}}(x)$.

We introduce three definitions that will be useful in the analysis. First, define the **convex hull** $t(\cdot) = \operatorname{conv} t_{\mathcal{X}}(\cdot)$ of the function $t_{\mathcal{X}}(\cdot)$ to be the greatest convex function majorized by the function (Rockafellar, 1970, page 36):

$$t(x) \equiv \inf\{t \,|(x,t) \in \operatorname{conv} \operatorname{epi} t_{\mathcal{X}}(\cdot)\},$$

where $\operatorname{conv} \operatorname{epi} t_{\mathcal{X}}(\cdot)$ is the convex hull of the epigraph of $t_{\mathcal{X}}(\cdot)$.[4] By definition, the convex hull of a function on $\mathcal{X}$ is a convex function with domain $\bar{\mathcal{X}}$ given by the convex hull of $\mathcal{X}$.

Second, define the **subdifferential** $\partial t(x)$ of a convex function $t : \bar{\mathcal{X}} \to \mathbb{R}$ evaluated at x to be the set of all **subgradients** $v \in \mathbb{R}^J$ such that for all $x' \in \bar{\mathcal{X}}$, $t(x') \geq t(x) + v \cdot (x' - x)$. The subdifferential of a convex function is a nonempty convex set at every point in the interior of its domain. Further, if t is differentiable at x, then the subdifferential is a singleton containing only the gradient of t evaluated at x, $\nabla t(x)$.

Third, define the **convex conjugate** t^* of t by

$$t^*(v) = \sup_{x \in \bar{\mathcal{X}}} \{v \cdot x - t(x)\},$$

the highest payoff available to a type v under the convexified transfer function. Subgradients

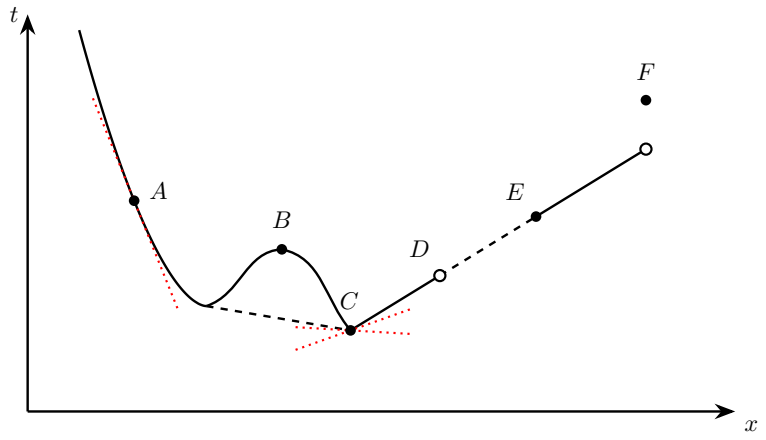
[3] Since $\mathcal{X}$ is the image of $\mathcal{A}$ under $x_{\mathcal{A}}$, the set over which the infimum is taken is nonempty for every $x \in \mathcal{X}$, so $t_{\mathcal{X}}(x) < \infty$. Since $t_{\mathcal{A}}(\cdot)$ is bounded below, $t_{\mathcal{X}}(x) > -\infty$.

[4] The epigraph of a function f is the set of points (x, y) such that $y \geq f(x)$.

of t and t^* are inverses of one another. That is, $x \in \partial t^*(v)$ if and only if $v \in \partial t(x)$ (Rockafellar, 1970, Theorem 23.5). We refer to this as the **inverse-subgradient relation**. It implies that $\partial t^*(v)$ is the set of outcomes in $\bar{\mathcal{X}}$ that maximize $v \cdot x - t(x)$.

Figure 1 illustrates the function $t_{\mathcal{X}}(\cdot)$, its convex hull $t(\cdot)$, and subgradients of $t(\cdot)$ at certain points. The horizontal axis denotes the potentially high-dimensional outcome space $\mathcal{X}$. The solid curve represents $t_{\mathcal{X}}(\cdot)$ and the dashed curve represents the parts where $t(\cdot)$ differs from $t_{\mathcal{X}}(\cdot)$. Specific pairs of x and t are labeled A through F, with coordinates $(x_A, t_A), \ldots, (x_F, t_F)$. The graph of $t_{\mathcal{X}}(\cdot)$ contains all the points achievable by some action in $\mathcal{A}$. Thus, this graph contains A, B, C, E and F, but not the hollow points. The graph of $t(\cdot)$ does not contain the dominated point B but contains D and all points in the segment $D - E$ which can be achieved, in expectation, by randomizing between C and E, but not by choosing any particular action in $\mathcal{A}$. The function $t(\cdot)$ can exhibit upward discontinuities. As x approaches x_F, $t(x)$ approaches the hollow point below F. However, at x_F, $t(x_F)$ jumps upwards to point F.

Figure 1: Outcomes, Convex Hulls and Subgradients



Notes: The solid curve is the graph of $t_{\mathcal{X}}(x)$ and the dashed curve is the graph of $t(x)$ for points where $t(x) < t_{\mathcal{X}}(x)$. The points labeled A through F represent outcome–transfer pairs. When a pair is represented by a solid circle, it belongs to the graph of $t_{\mathcal{X}}(x)$ and results from an action $a \in \mathcal{A}$. A pair represented by a hollow circle does not belong to the graph of $t_{\mathcal{X}}(x)$. Such points do not result from an action in the set $\mathcal{A}$.

Since $t_{\mathcal{X}}(\cdot)$ and $t(\cdot)$ are differentiable at x_A, the only element of the subdifferential is the slope of the dotted line through A. In contrast, these functions are not differentiable at C. The subdifferential of $t(\cdot)$ consists of the slopes of all lines that are everywhere below the dashed and solid lines. The only element of the subdifferential of $t(\cdot)$ of D and E is the slope of the line through the points $C - E$.

We make the following assumption to ensure that the optimization problem faced by each agent and their resulting choices are well-behaved:

Assumption 1. *(i) The function $t_{\mathcal{X}}(\cdot)$ is lower semicontinuous.*

(ii) The set $\mathcal{X}$ is nonempty and compact.

(iii) The support of the observed choices X conditional on $V = v$ is a subset of $\arg\max_{x \in \mathcal{X}} v \cdot x - t_{\mathcal{X}}(x)$.

Parts (i) and (ii) above imply, by the extreme value theorem, that the solution to the problem $\max_{x \in \mathcal{X}} v \cdot x - t_{\mathcal{X}}(x)$ for each $v \in \mathbb{R}^J$ is attained for some element of $\mathcal{X}$. Part (i) rules out upward discontinuities in $t_{\mathcal{X}}(x)$ as in the hollow point below F in figure 1 and part (ii) rules out the hollow point D. Compactness of $\mathcal{X}$ is either commonly satisfied or a weak restriction in the models discussed in section 3 below. For instance, in first-price auctions, $\mathcal{X}$ is the unit interval; in the hedonic demand models, it restricts the characteristic space to be compact. Part (iii) assumes that every outcome in the support of $X|V = v$ is optimal for each value of v. The maximum is attained by Assumptions 1(i) and 1(ii).

2.2 Noncooperative Games

Although we began with the single-agent case, our framework also allows analysis of a large class of simultaneous move games. Let a_{-i} and v_{-i} respectively denote the actions and values of agents other than i. Let $\tilde{x}_i(a_i, a_{-i})$ be the outcome and $\tilde{t}_i(a_i, a_{-i})$ be the transfer for agent i as a function of the action profile (a_i, a_{-i}). If the outcome is random conditional on the action profile, then interpret $\tilde{x}_i(a_i, a_{-i})$ and $\tilde{t}_i(a_i, a_{-i})$ as the corresponding expected values.

Assume that agent i's payoff playing a_i when the other agents play a_{-i} is given by

$$v_i \cdot \tilde{x}_i(a_i, a_{-i}) - \tilde{t}_i(a_i, a_{-i}).$$

Let $\mathcal{J}_i$ denote agent i's information set, with $v_i \in \mathcal{J}_i$. The expected utility from playing a_i is given by

$$v_i \cdot \mathbb{E}\left[\tilde{x}_i(a_i, a_{-i}) \middle| a_i; \mathcal{J}_i\right] - \mathbb{E}\left[\tilde{t}_i(a_i, a_{-i}) \middle| a_i; \mathcal{J}_i\right],$$

where expectations are taken with respect to the distribution of actions a_{-i} of the other players given the information set $\mathcal{J}_i$. Uncertainty in this model can arise either because i expects its opponents to play a mixed strategy, because the types v_{-i} are private information from the perspective of agent i, or both.

This model fits our framework if each agent is best responding to the distribution of actions played by others. Specifically, set

$$x_{\mathcal{A}}(a_i; \mathcal{J}_i) = \mathbb{E}\left[\tilde{x}_i(a_i, a_{-i}) \middle| a_i; \mathcal{J}_i\right]$$
$$t_{\mathcal{A}}(a_i; \mathcal{J}_i) = \mathbb{E}\left[\tilde{t}_i(a_i, a_{-i}) \middle| a_i; \mathcal{J}_i\right].$$

When analyzing games, we will assume that the two functions above are known to the analyst. A sufficient condition is that the outcome and the distribution of actions a_{-i} that agent i expects, $a_{-i}|\mathcal{J}_i$, is known or can be identified.

For example, consider the case in which agents' types are private information, each drawn independently from a distribution with CDF F_V. In this case, if the strategy profile $\sigma^*(v)$ constitutes a Bayesian Nash Equilibrium, then

$$x_{\mathcal{A}}(a_i) = \int \tilde{x}_i\left(a_i, \sigma^*_{-i}(v_{-i})\right) \mathrm{d}F_{V_{-i}} \tag{1}$$
$$t_{\mathcal{A}}(a_i) = \int \tilde{t}_i\left(a_i, \sigma^*_{-i}(v_{-i})\right) \mathrm{d}F_{V_{-i}}, \tag{2}$$

where the conditioning on the information set is dropped because it is irrelevant. Since play is described by a Bayesian Nash Equilibrium, agents have correct beliefs about the distribution of a_{-i}. The distribution of a_{-i} that each agent expects is identified from observation of the actions of all agents in independent and identically distributed instances of the game.

2.3 Observables

We follow the convention that upper-case letters denote random variables while lower case letters indicate specific values of the corresponding random variable.

Consider a dataset in which the analyst has access to a large sample of observations indexed by i from conditionally independent choices given a set of covariates Z. For each observation i, the analyst observes the covariates z_i and the chosen outcome x_i.

In addition, we assume that the analyst knows or observes the feasible outcomes $\mathcal{X}$ and the function $t_{\mathcal{X}}(\cdot)$. Therefore, we assume that $\mathcal{X}$ and $t_{\mathcal{X}}(\cdot)$ are either identified from the data or known from institutional details.

As discussed above, these outcomes and payoffs are generated by choosing actions $a \in \mathcal{A}$ with corresponding outcomes and transfers. Therefore, an alternative to observing x_i and having knowledge of $t_{\mathcal{X}}(\cdot)$ is that the actions a_i are observed and the functions $x_{\mathcal{A}}(\cdot)$ and

$t_{\mathcal{A}}(\cdot)$ are known. The rest of the analysis treats X and $t_{\mathcal{X}}(\cdot)$ as observable although all of our conclusions carry over to the alternative case by setting $X = x_{\mathcal{A}}(A)$. Thus, the distribution of X conditional on Z is identified directly from the observed data. Throughout, we write

$$P_{X|z}(\mathcal{X}') := \Pr(X \in \mathcal{X}' \mid Z = z) \tag{3}$$

for the probability that the outcome lies in a Borel set $\mathcal{X}' \subseteq \mathcal{X}$ conditional on $Z = z$.

We can also allow for $\mathcal{X}$ and the function $t_{\mathcal{X}}(x)$ to vary as long as the sample includes many observations for a given pair $(\mathcal{X}, t_{\mathcal{X}}(\cdot))$. We omit variation in this object for notational simplicity because our analysis will be conditional on fixed $\mathcal{X}$ and $t_{\mathcal{X}}(x)$.

In what follows, the conditioning on specific values of Z will be dropped from the notation except when we explicitly put together choices of agents with different values of the observables.

2.4 Identification

A model $(\mathbb{F}, \mathbb{S})$ is defined by a collection $\mathbb{F}$ of joint distributions of V and Z, $F_{V,Z}$, and a collection $\mathbb{S}$ of maps $\phi : \mathbb{F} \to \mathbb{H}$, where $\mathbb{H}$ is the set of all joint distributions of X and Z.

The collection $\mathbb{S}$ is generated by Assumption 1(iii) and a tie-breaking rule among optimal outcomes.

Definition 1. A **selection kernel** $K_{X^*|V,Z}$ is a conditional distribution of outcomes given (V, Z) that satisfies Assumption 1(iii). Let $\mathbb{K}$ be the set of selection kernels.

The support of each selection kernel $K \in \mathbb{K}$ conditional on $V = v$ and $Z = z$ is contained in the set of outcomes that are optimal for type v. Each $K \in \mathbb{K}$ induces a map $\phi_K : \mathbb{F} \to \mathbb{H}$. For every $F_{V,Z} \in \mathbb{F}$, $\phi_K(F_{V,Z})$ has marginal distribution F_Z on Z and conditional choice probabilities

$$P^K_{X|z}(\mathcal{X}') = \int K_{X^*|V=v,Z=z}(\mathcal{X}') \, \mathrm{d}F_{V|Z=z}(v) \tag{4}$$

for all z and all Borel $\mathcal{X}' \subseteq \mathcal{X}$. Thus, $\phi_K(F_{V,Z})$ is the distribution of observables produced when the selection kernel K maps $F_{V,Z}$ into conditional choice probabilities $P^K_{X|z}$. We consider the models generated by $\mathbb{K}$: $\mathbb{S} = \{\phi_K : K \in \mathbb{K}\}$.

We follow the standard identification and falsifiability definitions in the literature (e.g., Athey and Haile, 2002; Matzkin, 2007). A model is identified if the joint distribution of the model's primitives is uniquely determined by the joint distribution of observables. It is falsifiable if

some distributions of observables are inconsistent with it. Just as identification of a model is necessary but not sufficient for the existence of a consistent estimator, falsifiability is necessary but not sufficient for the existence of a valid statistical test (Berry and Haile, 2018). Both estimation and inference require additional statistical analysis. We leave such analyses for future research.

We define the identified set $\mathcal{F}_H$ for observables H following this literature:

Definition 2. The identified set for a distribution of observables $H \in \mathbb{H}$ is

$$\mathcal{F}_H = \{F_{V,Z} \in \mathbb{F} : \exists \phi \in \mathbb{S} \text{ such that } \phi(F_{V,Z}) = H\}. \tag{5}$$

3 Examples

3.1 Single-Agent Problems

Example 1. Hedonic Demand Models. Consider the hedonic demand model in which consumer i's indirect utility from purchasing good $k \in \{1, \ldots, K\}$ is given by

$$\sum_j x_{kj}\beta_{ij} - p(x_k),$$

where $p(\cdot)$ is the pricing function, x_{kj} denotes the $j-$th characteristic of product k, and β_{ij} denotes the random coefficients. This model fits our framework with $v_i = (\beta_{i1}, \ldots, \beta_{iJ})$ and $t(\cdot) = p(\cdot)$ and corresponds to the hedonic demand model proposed in Gorman (1980) and Lancaster (1966) with the additional restriction that preferences are linear in characteristics and prices. Rosen (1974) proposed estimating such models by first estimating $p(\cdot)$ and then using data on purchases to recover β_{ij}, the marginal willingness to pay for x_{kj}. Bajari and Benkard (2005) incorporate price endogeneity into this framework by including an unobserved quality index $\xi_k \in \mathbb{R}$ into x_k and show how to identify the pricing function and the unobserved quality of each good in a first step. Our analysis takes this first step as given and focuses on identifying the marginal willingness to pay.

Example 2. Multinomial Choice with Exogenous Characteristics. Consider a discrete choice model (see McFadden, 1973; Train, 2009). An outcome $x \in \mathcal{X} = \left\{x \in \{0,1\}^J : \sum_j x_j \leq 1\right\}$ denotes which option is chosen by a consumer. Let $t(x)$ denote the price of option x and let v_i denote the vector of utilities for the other attributes. The consumer's utility from picking

any given $x \in \mathcal{X}$ is given by the form assumed in our model:

$$v_i \cdot x - t(x).$$

A large literature (e.g., Berry et al., 1995) focuses on solving the price endogeneity problem. This problem is particularly relevant for the consumer choice context when certain product attributes are unobserved. We do not address endogeneity of this form in our analysis, assuming that the researcher is able to find a solution in a prior step or it is reasonable that endogeneity is not a concern. Thus, our results are relevant to choices in which unobserved product attributes can be controlled for by other means or if there is consumer-level price variation within the market (e.g., Tebaldi et al., 2019).

3.2 Mechanisms and Trading Games

Example 3. Single and Multi-Unit, Independent Private Value (IPV) Auctions. Our model embeds standard IPV auctions that award the object to the highest bidder. In the single-unit case, the action $a \in \mathbb{R}$ corresponds to a bid, v_i is agent i's value for the object, $x_{\mathcal{A}}(a) \in [0,1]$ denotes the probability of winning with a bid a, and $t_{\mathcal{A}}(a)$ denotes the expected payment. We can accommodate both first-price and all-pay auctions, amongst others. For example, in the first-price auction analyzed by Guerre et al. (2000), each agent chooses a bid a to maximize $(v-a)G(a)$, where $G(a)$ is the cumulative distribution function of the highest bid of the remaining bidders. In a Bayesian Nash Equilibrium, bidders have consistent beliefs about the bid distribution of opposing bidders and data from many independent and identical auctions identifies $G(a)$. The bid a parameterizes $x_{\mathcal{A}}(a) = G(a)$ and $t_{\mathcal{A}}(a) = aG(a)$. The multi-unit case (Cassola et al., 2013, e.g.) can also be incorporated by setting $x_{\mathcal{A}}(a_i)$ to be a vector with $k-$th element equal to the win probability of the $k-$th unit, and $t_(a_i)$ to the expected payment depending on the auction format (uniform or discriminatory).

We assume that all bids are observed. The identification of these auctions under fewer data requirements has been analyzed in Athey and Haile (2002).

Example 4. Mechanisms with Private Information, Independent Types, and Quasi-linear Utilities. Our model can nest mechanisms more generally. Consider a mechanism described by the allocation rule $\tilde{x}_i(a_i, a_{-i})$ and transfer function $\tilde{t}_i(a_i, a_{-i})$, where the set of actions $\mathcal{A}$ coincides with the set of messages an agent can send and each agent's valuation is their private type. In incentive-compatible direct mechanisms, agents will truthfully reveal their types and the distribution of a_{-i} will coincide with the distribution of valuations. In

indirect mechanisms, revealed preference arguments link valuations and messages. Moreover, the resulting allocation may involve ironing or pooling types. Pooling is particularly important if the space of messages is discrete, for example if there are minimum bid increments. An example with discrete messages includes scoring auctions in which bids may include binary services in addition to a continuous monetary amount (see Aspelund and Russo, 2023, for example). Our approach will allow for both cases.

Example 5. Oligopoly Pricing. The identification of marginal costs in canonical models (e.g., Berry et al., 1995) is based on first identifying demand using cost shifters (see Berry and Haile, 2014), then assuming a Nash Equilibrium in simultaneous move price setting game. To map this exercise to our notation, let a_i be the price chosen by firm i, v_i be the negative of the firm's marginal cost, $\tilde{x}(a_i, a_{-i})$ be the quantity sold by firm i, and $\tilde{t}(a_i, a_{-i})$ be the negative of the firm's revenue. Therefore, $x_{\mathcal{A}}(a_i; \mathcal{J}_i)$ and $t_{\mathcal{A}}(a_i; \mathcal{J}_i)$ are the expected quantities and the negative of expected revenues, respectively. Profit maximization implies that the firm maximizes

$$v_i \cdot x_{\mathcal{A}}(a_i; \mathcal{J}_i) - t_{\mathcal{A}}(a_i; \mathcal{J}_i).$$

We can allow for information sets that may or may not include the prices set by other firms by varying $\mathcal{J}_i$.

Instead of a price setting game, it is also straightforward to fit a quantity setting game into our model by interpreting a as quantities and setting $x_i(a_i; \mathcal{J}_i) = a_i$. The function $t_i(a_i; \mathcal{J}_i)$ still denotes the expected negative revenue.

Example 6. School Choice. Our model can also be applied to settings where there are no monetary transfers. For instance, Agarwal and Somaini (2018) consider a school assignment mechanism in which agents can submit rank order lists indexed by a.[5] Assume that preferences are private information and each student knows the distribution from which the preferences of other students are drawn. In equilibrium, a student submitting the list a is assigned to one of J schools with probability vector $x_{\mathcal{A}}(a) \in \Delta^J \subseteq \mathbb{R}^J$, where $x_{\mathcal{A}}(a)$ is derived as in section 2.2. Let $d_i = (d_{i1}, \ldots, d_{iJ})$ be the vector of distances of each school from student i, and let u_i denote the vector of indirect utilities from assignment into each school, net of distance. If preferences are linear in distance, then the expected utility from submitting list a is given by

$$(u_i - d_i) \cdot x_{\mathcal{A}}(a).$$

This model fits our framework by setting $t_{\mathcal{A}}(a; \mathcal{J}_i) = 0$ and $v_i = u_i - d_i$.

[5] We fix the priority type of the student and drop it from the notation for simplicity.

Example 7. Trading Games. Larsen and Zhang (2018) consider a trading game in which an agent can take a sequence of actions $a = (a_1, \ldots, a_M)$. Following this sequence, the agent has a probability $x_{\mathcal{A}}(a) \in [0,1]$ of engaging in a transaction and paying an expected (possibly negative) transfer $t_{\mathcal{A}}(a)$. The value of the trade for agent i is given by $v_i \in [\underline{v}_i, \bar{v}_i] \subseteq \mathbb{R}$. Therefore, the expected utility from the action a sequence is

$$v_i \cdot x_{\mathcal{A}}(a) - t_{\mathcal{A}}(a),$$

where the actions of other agents are integrated over as in section 2.2.

4 Revealed Preferences

4.1 Rationalizable Actions

Our first result derives a testable implication of the model and characterizes the outcomes x that are optimal for some $v \in \mathbb{R}^J$. For a convex set C, $\mathrm{ri}(C)$ and $\mathrm{relbd}(C)$ denote the interior and boundary of C relative to its affine hull (Rockafellar, 1970, Section 6).

Proposition 1. *(i) The outcome $x \in \mathcal{X}$ is optimal for type $v \in \mathbb{R}^J$ if and only if $t_{\mathcal{X}}(x) = t(x)$ and $v \in \partial t(x)$. (ii) If $t_{\mathcal{X}}(\cdot)$ is convex and lower-semicontinuous, then for every $x \in \mathcal{X} \cap \mathrm{ri}\left(\bar{\mathcal{X}}\right)$ there exists $v \in \mathbb{R}^J$ such that x is an optimal choice.*

Proof. Part (i). If: Let $v \in \partial t(x)$ and $t_{\mathcal{X}}(x) = t(x)$. Since $v \in \partial t(x)$, for all x' in the convex hull of $\mathcal{X}$, $v \cdot (x' - x) \leq t(x') - t(x)$. Observe that $t_{\mathcal{X}}(x) = t(x)$ and $t(x') \leq t_{\mathcal{X}}(x')$ for all $x' \in \mathcal{X}$. Therefore, we have that $v \cdot (x' - x) \leq t_{\mathcal{X}}(x') - t_{\mathcal{X}}(x)$ as required.

Only if: Suppose that $x \in \mathcal{X}$ is optimal for v. Towards a contradiction suppose that $t(x) < t_{\mathcal{X}}(x)$. Since $t(\cdot)$ is the convex hull of $t_{\mathcal{X}}(\cdot)$, there exist $x_j \in \mathcal{X}$ and weights $\alpha_j \geq 0$ such that $\sum \alpha_j = 1$, $\sum_j \alpha_j x_j = x$ and $\sum_j \alpha_j t_{\mathcal{X}}(x_j) = t(x)$. Therefore,

$$v \cdot x - t_{\mathcal{X}}(x) < \sum_j \alpha_j \left(v \cdot x_j - t_{\mathcal{X}}(x_j)\right).$$

Hence, it must be that there exists x_j such that $v \cdot x - t_{\mathcal{X}}(x) < v \cdot x_j - t_{\mathcal{X}}(x_j)$. This contradicts the assumption that x is optimal. Thus, $t(x) \geq t_{\mathcal{X}}(x)$, and by definition of $t(\cdot)$, $t(x) \leq t_{\mathcal{X}}(x)$. Therefore, $t_{\mathcal{X}}(x) = t(x)$.

To show that $v \in \partial t(x)$, assume towards another contradiction that $v \notin \partial t(x)$. That is, assume there exists x' in the convex hull of $\mathcal{X}$ such that $v \cdot x - t(x) < v \cdot x' - t(x')$. Hence, there

exist $x_j \in \mathcal{X}$ and weights $\alpha_j \geq 0$ such that $\sum \alpha_j = 1$ and $\sum_j \alpha_j \left(v \cdot x_j - t_{\mathcal{X}}\left(x_j\right)\right) = v \cdot x' - t\left(x'\right)$. Since $t_{\mathcal{X}}\left(x\right) = t\left(x\right)$, it must be that there exists x_j such that $v \cdot x - t_{\mathcal{X}}\left(x\right) < v \cdot x_j - t_{\mathcal{X}}\left(x_j\right)$. This contradicts the assumption that x is optimal.

Part (ii). Under the maintained assumptions, $t_{\mathcal{X}}\left(x\right) = t\left(x\right)$ for all $x \in \mathcal{X}$ and $\partial t_{\mathcal{X}}\left(x\right) = \partial t\left(x\right)$. Moreover, t is a proper convex function, so $\partial t\left(x\right)$ is nonempty at every point of $\text{ri}\left(\bar{\mathcal{X}}\right)$. For $x \in \mathcal{X} \cap \text{ri}\left(\bar{\mathcal{X}}\right)$, Part (i) implies the result. □

Part (i) shows that linearity of payoffs in the model has testable implications. Specifically, outcomes x that result in payments larger than $t\left(x\right)$ are dominated. Figure 1 presents an illustration. Consider the point A and an agent with type v_A given by the slope of the tangent through A. Lines parallel to this tangent are the indifference curves for an agent with type v_A, with utility increasing as t decreases. Therefore, the unique optimizer for this agent is x_A. In contrast, points in the neighborhood of x_B are not optimal. An agent would rather pick either x_C or an outcome close to x_A. The outcome x_C is also optimal, but for more than one type. The slopes of the two dotted lines touching C, and convex combinations of them, are the types for which C is optimal.

Part (ii) shows that if the cost function $t_{\mathcal{X}}\left(\cdot\right)$ is convex and lower semicontinuous, every $x \in \mathcal{X} \cap \text{ri}\left(\bar{\mathcal{X}}\right)$ can be rationalized. By contrast, some points of $\mathcal{X}$ on $\text{relbd}\left(\bar{\mathcal{X}}\right)$ cannot be rationalized. For example, if $t_{\mathcal{X}}\left(x_n\right) \to \infty$ along a sequence $x_n \in \mathcal{X}$ with $x_n \to x$, then no type makes x optimal. Convexity implies that $t_{\mathcal{X}}\left(x\right) = t\left(x\right)$ for all $x \in \mathcal{X}$, ruling out points like B in figure 1. Lower semicontinuity rules out upward discontinuities like the one at F in figure 1. The outcome $x_D \notin \mathcal{X}$ and therefore it does not belong to the domain of $t_{\mathcal{X}}\left(\cdot\right)$ although it belongs to the domain of $t\left(\cdot\right)$. The outcome x_F is not optimal for any type because an agent can achieve an arbitrarily close outcome to x_F while discretely reducing the payment. Proposition 1 implies that x_F is not optimal because $\partial t\left(x_F\right)$ is empty. Together, convexity and lower semicontinuity imply that all solutions to the problem $\max_{x \in \mathcal{X}} v \cdot x - t_{\mathcal{X}}\left(x\right)$ also solve the problem $\max_{x \in \bar{\mathcal{X}}} v \cdot x - t\left(x\right)$.

The assumptions of convexity and semicontinuity are commonly imposed in the literature on identification. In first-price auctions (example 3), it is often assumed that $a + \frac{G(a)}{\partial G(a)/\partial a}$ is increasing in a, where a is a bid and $G\left(a\right)$ is the cumulative distribution function of the highest competitor bid (see Assumption C2 in Guerre et al., 2000, for example). This assumption is identical to assuming that $t\left(x\right)$ is convex if it is twice differentiable. To see this, note that the first-derivative of $t\left(x\right) = xG^{-1}\left(x\right)$ is equal to

$$\nabla t\left(x\right) = G^{-1}\left(x\right) + \frac{x}{\partial G\left(G^{-1}\left(x\right)\right)/\partial a}$$

and substitute $x = G(a)$. Similarly, in the hedonic demand model, if the price function, $t(x)$, were nonconvex, then certain choices would be suboptimal under the linear random coefficients structure.

It is useful to define two nested subsets of $\mathcal{X}$. By Proposition 1(i), an outcome x is optimal for some type if and only if $t_{\mathcal{X}}(x) = t(x)$ and $\partial t(x) \neq \emptyset$. We refer to these outcomes as the rationalizable actions and denote them by $\mathcal{X}^r$. We further denote by $\mathcal{X}^\dagger$ the set of outcomes x for which $(x, t_{\mathcal{X}}(x))$ is an exposed point of the convex hull of the epigraph of $t_{\mathcal{X}}$, i.e., a point at which some supporting hyperplane meets the epigraph only there. The sets $\mathcal{X}^r$ and $\mathcal{X}^\dagger$ are Borel (Choquet et al., 1966). Moreover, we have that $\mathcal{X}^\dagger \subseteq \mathcal{X}^r \subseteq \mathcal{X}$.

Corollary 1. *(i) Any distribution of observables generated by the model satisfies $P_{X|z}(\mathcal{X}^r) = 1$. Consequently, if $\mathbb{F}$ is unrestricted over distributions of (V, Z), the model is falsifiable if and only if $\mathcal{X}^r \subsetneq \mathcal{X}$. (ii) Let $\mathbb{F}_d$ be the set of distributions $F_{V,Z} \in \mathbb{F}$ that admit a density for every z. Every distribution of observables generated by $(\mathbb{F}_d, \mathbb{S})$ satisfies $P_{X|z}\left(\mathcal{X}^\dagger\right) = 1$; consequently $(\mathbb{F}_d, \mathbb{S})$ is falsifiable whenever $\mathcal{X}^\dagger \subsetneq \mathcal{X}$.*

Proof. Part (i) follows from Proposition 1(i): an outcome outside $\mathcal{X}^r$ is optimal for no type and cannot be rationalized, whereas if $\mathcal{X}^r = \mathcal{X}$ every distribution of X is generated by some type distribution. For part (ii), the conjugate t^* is differentiable Lebesgue-almost everywhere, and at each point of differentiability the set of optimizers $\partial t^*(v)$ is a singleton. If $F_{V|Z=z}$ admits a density, the type has a unique optimal action with probability one, and that action belongs to $\mathcal{X}^\dagger$ by definition. Therefore every distribution of observables generated by $(\mathbb{F}_d, \mathbb{S})$ satisfies $P_{X|z}\left(\mathcal{X}^\dagger\right) = 1$; when $\mathcal{X}^\dagger \subsetneq \mathcal{X}$, $(\mathbb{F}_d, \mathbb{S})$ is falsifiable. □

4.2 Characterization of the Identified Set

This subsection characterizes the identified set of type distributions $F_{V|Z}$ given data on the conditional distribution of X given Z. Beyond Assumption 1, we impose no further restrictions on $\mathcal{X}$ or $t(\cdot)$. For example, the choice set may be discrete, implying that $\mathcal{X}$ is not a connected subset of $\mathbb{R}^J$ and $\partial t(x)$ is not a singleton. The following subsections will further specialize the results to the cases in which $t(\cdot)$ is differentiable.

We characterize the identified set in two equivalent ways. One indexes the revealed-preference restrictions by sets of outcomes: a candidate type distribution must put enough mass on types that can choose each observed event. The other indexes the same restrictions by sets of types: no set of types can be more probable than the outcomes that are optimal for types in that set. Both turn out to be sharp.

The first is based on the restrictions that optimality places on the choice probabilities. By Proposition 1 and Corollary 1, a distribution $F_{V,Z}$ rationalizes the observed choice probabilities only if $P_{X|z}(\mathcal{X}^r) = 1$ and, for every subset $\mathcal{X}' \subseteq \mathcal{X}$ of outcomes, the probability of choosing an element of $\mathcal{X}'$ does not exceed the mass of types for which some element of $\mathcal{X}'$ is optimal:

$$P_{X|z}(\mathcal{X}') \leq \int 1\{v \in \cup_{x\in\mathcal{X}'}\partial t(x)\}\, \mathrm{d}F_{V|Z=z}. \tag{6}$$

For $\mathcal{X}' \subseteq \mathcal{X}^r$, the right-hand side is the capacity functional of the original optimal-choice correspondence (Molchanov, 2017). Let $\mathcal{R}_H$ denote the set of distributions $F_{V,Z} \in \mathbb{F}$ such that $P_{X|z}(\mathcal{X}^r) = 1$ and these inequalities hold for all z and all $\mathcal{X}' \subseteq \mathcal{X}$.

The second restricts the type distribution directly. For every Borel set $\mathcal{V}$ of types, the mass of types in $\mathcal{V}$ does not exceed the probability of observing an outcome that is optimal for some type in $\mathcal{V}$. Let $\mathcal{O}_H$ denote the set of distributions $F_{V,Z} \in \mathbb{F}$ such that $P_{X|z}(\mathcal{X}^r) = 1$ and these containment inequalities

$$\int_{\mathcal{V}} \mathrm{d}F_{V|Z=z} \leq P_{X|z}(\cup_{v\in\mathcal{V}}\partial t^*(v)) \tag{7}$$

hold for all z and all Borel type sets $\mathcal{V}$. The right-hand side of (7) is the containment functional of the same correspondence, the probability that it is contained in a given set, just as the right-hand side of (6) is the probability that it intersects one.

Because any distribution in the identified set must rationalize the data, $\mathcal{F}_H \subseteq \mathcal{R}_H$. The next result shows that this necessary condition is also sufficient, so that these inequalities are sharp and give three equivalent descriptions of the identified set.

Proposition 2. *Suppose Assumption 1 is satisfied. The identified set is characterized equivalently by the capacity inequalities (6) and the containment inequalities (7): $\mathcal{F}_H = \mathcal{R}_H = \mathcal{O}_H$.*

Proof. If $P_{X|z}(\mathcal{X}^r) < 1$ for some z, then $\mathcal{R}_H = \mathcal{O}_H = \emptyset$ and $\mathcal{F}_H = \emptyset$. Now suppose $P_{X|z}(\mathcal{X}^r) = 1$ for all z. Any distribution in $\mathcal{F}_H$ rationalizes the data, so $\mathcal{F}_H \subseteq \mathcal{R}_H$. Conversely, if $F_{V,Z} \in \mathcal{R}_H$, the appendix constructs $K_{X^*|V,Z} \in \mathbb{K}$ reproducing the choice probabilities, so $\mathcal{R}_H \subseteq \mathcal{F}_H$. Lemma 3 in the appendix gives $\mathcal{O}_H = \mathcal{R}_H$, hence $\mathcal{F}_H = \mathcal{R}_H = \mathcal{O}_H$. □

Proposition 2 therefore gives three equivalent descriptions of the same set. Two of them are the inequality systems defining $\mathcal{R}_H$ and $\mathcal{O}_H$. The third is $\mathcal{F}_H$ itself, which describes the set through the selection kernels of Definition 1. Each description serves a different purpose. The capacity inequalities (6) are the revealed-preference restrictions that optimality places on the type distribution. They are the model's testable content (Corollary 1) and the moment

inequalities on which estimation and inference are based. The selection representation (4) is generative. It says that the data can be reproduced by some tie-breaking among optimal actions, which is the sole source of the model's incompleteness, and when $F_{V|Z}$ admits a density, so that ties are negligible, it yields a likelihood (Corollary 2). The containment inequalities (7) constrain the type distribution directly, which makes them the natural tool for bounding features of $F_{V|Z}$ such as payoffs, margins and counterfactuals in the application. They also deliver point identification in special cases. When t is differentiable each outcome is optimal for a unique type, so applying (7) to any type set and its complement forces both inequalities to bind and pins down $F_{V,Z}$ (Corollary 3).

Corollary 2. *Suppose Assumption 1 holds and $F_{V|Z=z} \in \mathcal{F}_H$ admits a density. Then for all z and all Borel $\mathcal{X}' \subseteq \mathcal{X}$,*

$$P_{X|z}\left(\mathcal{X}'\right) = \int 1\left\{v \in \cup_{x \in \mathcal{X}'} \partial t\left(x\right)\right\} \mathrm{d}F_{V|Z=z}.$$

In particular, the choice probabilities are determined by $F_{V|Z}$, so the model delivers a likelihood for the observed choices.

Proof. By Corollary 1(ii), the type has a unique optimal action with probability one, which we denote $\nabla t^*\left(v\right)$. Since $F_{V|Z=z} \in \mathcal{F}_H$, Proposition 2 yields a selection kernel $K_{X^*|V,Z} \in \mathbb{K}$ reproducing the choice probabilities. At almost every v the optimal set is the singleton $\{\nabla t^*\left(v\right)\}$, whose unique element belongs to $\mathcal{X}^\dagger$ and is therefore also the unique optimizer of the original problem, so $K_{X^*|V=v}$ is degenerate and $\int 1\left\{X^* \in \mathcal{X}'\right\} \mathrm{d}K_{X^*|V=v} = 1\left\{\nabla t^*\left(v\right) \in \mathcal{X}'\right\}$. For such v, $\nabla t^*\left(v\right) \in \mathcal{X}'$ if and only if $v \in \cup_{x \in \mathcal{X}'} \partial t\left(x\right)$, because the unique optimizer lies in $\mathcal{X}'$ exactly when $v \in \partial t\left(x\right)$ for some $x \in \mathcal{X}'$. Integrating over the full-measure set of such v gives the equality. □

In contrast to the prior literature on partially identified models, where incompleteness typically arises from equilibrium selection among multiple equilibria (Galichon and Henry, 2011; Beresteanu et al., 2011), incompleteness in our model arises from ties among optimal actions. The set of types for which multiple outcomes are optimal has Lebesgue measure zero, but may have positive probability under $F_{V|Z=z}$. Restricting the conditional type distributions to admit densities rules out these ties and yields a complete model. Even then the identified set need not be a singleton, because $\partial t\left(x\right)$ need not be a singleton.

The characterization yields an empirical analog that can be used as the basis of a two-step estimation strategy. The first step recovers the subgradients $\partial t\left(x\right)$, which may or may not

be singletons and which are the sets that enter the capacity inequalities (6). Computing the subgradients may first require estimating the function $t(x)$ in some contexts. The second step estimates bounds on the CDF $F_{V|Z}$ by imposing the containment inequalities (7) that define $\mathcal{O}_H$. A common example is auctions, a case that we further develop in our application in section 5.[6]

Another implication of the characterization is that considering only a subset of deviations weakens the link between types and outcomes. Each observed action is then consistent with a weakly larger set of types, and each type with a weakly larger set of outcomes. The right-hand sides of both (6) and (7) therefore weakly increase, so the inequalities are easier to satisfy and the identified set is weakly larger. This implication is expected because fewer necessary conditions for optimality are imposed. A less obvious question is whether one can still derive a likelihood function. This turns out to be true if $F_{V|Z}$ admits a density even if we impose only a subset of deviations (Heitjan and Rubin, 1991). This implication will be important in our analysis of combinatorial auctions.

The characterization in Proposition 2 is a generalization of the one in Larsen and Zhang (2018). Specifically, consider an agent, indexed by i, who picks actions that result in a probability of trade equal to x_i. Assume that the agent can also pick actions that instead yield either $x_i - \Delta_1$ or $x_i + \Delta_2$ for some $\Delta_1, \Delta_2 > 0$. Since x_i is optimal, it must be that

$$\frac{t(x_i) - t(x_i - \Delta_1)}{x_i - (x_i - \Delta_1)} \leq v_i \leq \frac{t(x_i) - t(x_i + \Delta_2)}{x_i - (x_i + \Delta_2)}.$$

Convexity of $t(\cdot)$ implies that the local deviations considered above are both necessary and sufficient for optimality.

Our result extends this logic to higher dimensions. As an example, consider a hedonic demand model with a discrete product set. In this context, our result identifies the set to which the vector representing consumers' willingness to pay for specific product characteristics belongs. When the support of actions, and consequently the support of x, is discrete, the set $\partial t(x)$ at some of these points in the support has positive volume in $\mathbb{R}^J$. However, if the action space includes both discrete and continuous elements, the set $\partial t(x)$ may neither be a singleton nor have positive volume in $\mathbb{R}^J$. Instead, it forms a convex set with an affine hull $\text{aff}(\partial t(x))$ whose dimension lies between 0 and J.

[6] Aspelund and Russo (2023) is a recent example that analyzes a scoring auction using this estimation procedure in a context that does not squarely fit into prior models.

4.3 Identification with Rich Choice Environments

Our next result considers the special case when $t(x)$ is differentiable. Identification of an individual agent's type is a corollary of Proposition 1 because differentiability at x_i makes $\partial t(x_i)$ a singleton, so v_i is identified. Identification of the full distribution $F_{V,Z}$ is a corollary of the characterization in Proposition 2: differentiability makes each action optimal for a unique type, so the containment inequalities (7) defining $\mathcal{O}_H = \mathcal{F}_H$, applied to any set of types and its complement, hold with equality and pin down $F_{V,Z}$.

Corollary 3. *Suppose Assumption 1 is satisfied and agent i chooses outcome $x_i \in \mathcal{X}$. If $t(\cdot)$ is differentiable at x_i, then v_i is identified. In particular, if $t(\cdot)$ is differentiable for all x in the support of X, then $F_{V,Z}$ is identified.*

In a rich choice environment, each observed choice is compatible with at most one payoff type. For this reason, we label such a choice environment as "rich."

It is instructive to apply the result to first-price auctions and to hedonic choice models. Recall our calculation of the derivative of $t(x) = xG^{-1}(x)$ in first-price auctions (example 3). This derivative, calculated at $x = G(b)$, is precisely the private value implied by the first-order condition of a bidder bidding b. In the higher-dimensional case of hedonic demand models, the marginal cost of increasing x_j is $\partial t(x)/\partial x_j$. Therefore, if consumer i chooses a product or a consumption bundle described by x_i, then optimality implies that $\nabla t(x_i)$ is equal to the vector of marginal willingness to pay for each of the components, v_i.

4.4 Identification with Preference Shifters

We now show that even without rich choice sets, variation in an observable shifter of preference, denoted z, can be used to point identify the model. We start with the case when z is a special regressor (Lewbel, 2000), that is, $v = u + z$, with $u \perp z$. Then, we generalize this result to the case when $v = u + g(z)$, where $g(\cdot)$ is a nonlinear function.

The rest of this section makes the following assumptions on the distribution of u and the support of outcomes:

Assumption 2. *The random variable u is independent of z, admits a density $f_U(u)$ and has a moment generating function.*

This assumption is equivalent to the existence of a constant $k > 0$ such that $\exp(k|u|) f_U(u)$ is Lebesgue integrable.[7] It requires that the tails of u decline sufficiently rapidly and is satisfied by most commonly used parametric forms, including multivariate normals and extreme-value distributions as well as finite mixtures of these distributions.

Assumption 3. *The convex hull of $\mathcal{X}$ has a nonempty interior.*

This assumption allows choice sets that are not rich in the sense defined above but requires their convex hull to be full-dimensional, so choices can be informative about every dimension of payoffs. In particular, if $\mathcal{X}$ contains finitely many options, $t(\cdot)$ is nondifferentiable at the exposed points $\mathcal{X}^\dagger \subseteq \mathcal{X}$. If valuations admit a density, however, $P_{X|z}\left(\mathcal{X}^\dagger\right) = 1$, implying that every choice made with positive probability has a nonsingleton $\partial t(x)$.

Theorem 1. *Suppose that $v = u + z$ and Assumptions 1, 2 and 3 are satisfied.*

(i) If Z has full support on $\mathbb{R}^J$, then $f_U(\cdot)$ is identified.

(ii) More generally, suppose that Z has full support on a linear subspace $L \subseteq \mathbb{R}^J$. If at every $x \in \bar{\mathcal{X}}$, $t(x)$ is differentiable in every direction orthogonal to L, then $f_U(\cdot)$ is identified.

Part (ii) trades variation in z against smoothness of t. Shifter variation identifies preferences along the directions Z spans, while differentiability identifies them along the remaining directions, where each observed choice reveals the type exactly. Part (i) is the extreme case in which Z spans all of $\mathbb{R}^J$ and no differentiability is required.

The argument proceeds in a few steps. By Lemma 1 in the appendix, there is a set $A \subseteq \mathcal{X}$ whose image $C := \cup_{x \in A} \partial t(x)$ has strictly positive Lebesgue measure and is contained in a translated simplicial cone.[8] Because Assumption 2 implies that $V|Z = z$ admits a density, Corollary 2 gives:

$$P_{X|z}(A) = \int 1\{u + z \in C\} f_U(u)\, \mathrm{d}u.$$

The observed choice probabilities are thus a linear transform of f_U, and identifying f_U reduces to showing this transform is injective. That is, $\int 1\{u + z \in C\} f(u)\, \mathrm{d}u = 0$ for all z implies $f(u) = 0$ almost everywhere, so that variation in z pins down the distribution of u. Because C is a subset of a simplicial cone, we can pick a vector λ in the interior of the cone's dual

[7] The moment generating function of u is defined as $M_U(t) = Ee^{tU} = \int e^{tu} f_U(u)\, du$. The moment generating function exists if the integration exists in a neighborhood of zero.

[8] A simplicial cone is the set of nonnegative linear combinations of J linearly independent vectors. A translated simplicial cone is such a set shifted by a fixed vector.

cone that is small enough so that $\exp(2\pi u \cdot \lambda) f(u)$ remains integrable. A Fourier transform then re-writes the injectivity condition as

$$\hat{\chi}_{C,\lambda}(\xi)\,\hat{f}_{\lambda}(\xi) = 0,$$

where $\hat{f}_{\lambda}$ and $\hat{\chi}_{C,\lambda}$ are the Fourier transforms of $\exp(2\pi u \cdot \lambda) f(u)$ and $1\{u \in C\}\exp(-2\pi u \cdot \lambda)$. If f were nonzero, $\hat{f}_{\lambda}$ would be nonzero on an open set, forcing $\hat{\chi}_{C,\lambda}$ to vanish there. Lemma 2 in the appendix rules this out. The function $\hat{\chi}_{C,\lambda}$ extends to a holomorphic function, so vanishing on an open set would force it to vanish everywhere, and it cannot vanish everywhere because C has strictly positive Lebesgue measure. Hence f_U is identified. Part (ii) follows by applying this argument within L. Differentiability in the directions orthogonal to L means that each observed choice reveals those components of v exactly, and conditioning on them leaves a problem in which Z has full support on L, where part (i) applies. Formal details are in appendix A.

A limitation of Theorem 1 is that it requires a linearly separable regressor z that is independent of u. Our next result relaxes this requirement by considering the model

$$v = u + g(z),$$

where $g : \mathbb{R}^J \to \mathbb{R}^J$ with the dimension of z, $d_z \geq J$. We use an argument inspired by Allen and Rehbeck (2017) to show that the function $g(\cdot)$ is identified under the following additional restrictions:

Assumption 4. *(i) $J \geq 2$ and the function $g(z)$ is given by $(g_1(z_1), \ldots, g_J(z_J))$. Moreover, each $g_j(\cdot)$ is continuously differentiable with a nonzero derivative at each point.*

(ii) For any l, $k \in \{1, \ldots, J\}$, the partial derivatives of each $\mathbb{E}[X_l|z]$ with respect to z_k exist, are continuous, and are nonzero for all z.

(iii) The support of Z is rectangular.

(iv) The expectation of U exists.

The main restrictions are in parts (i) and (ii). Part (i) assumes that each of the regressors z_j is component-specific and that there are at least two components. We will discuss the need for at least two components after presenting our main result. Extensions that allow for additional regressors, some of which are common, can be accommodated by following arguments in Allen and Rehbeck (2017). Part (ii) assumes that the expected value of the optimal x is smooth and the partial derivatives are nonzero. This assumption would be

violated if a component of g were not globally either a strict complement or substitute with a component of X. Since this quantity is observed in the data, this assumption is falsifiable.[9]

Parts (iii) and (iv) are technical regularity assumptions. Part (iii) allows us to move from knowing the derivatives of g to determining the function up to a location. Part (iv) is a weak condition that is implied, for example, by Assumption 2.

Proposition 3. *Suppose that Assumptions 1 and 4 are satisfied and $u \perp z$. Then, g is identified on its support up to the location and scale normalizations $g(z_0) = 0$ and $\frac{\partial}{\partial z_j} g_j(z_0) \in \{-1, 1\}$ for some j, respectively.*

We sketch the argument here and give formal details in appendix A. Define

$$W(z) = \mathbb{E}\left[\max_{x \in \mathcal{X}} \{x \cdot (U + g(z)) - t(x)\}\right].$$

The main observation is a subgradient-envelope argument showing that $W(z)$ is differentiable and that

$$\nabla W(z) = Dg(z)^\top \mathbb{E}[X|z].$$

Because g is component-specific, this gives

$$\frac{\partial W(z)}{\partial z_k} = \mathbb{E}[X_k|z]\, g_k'(z_k).$$

Differentiating once more with respect to z_l for $l \neq k$ and using the equality of mixed partial derivatives, $\frac{\partial^2 W(z)}{\partial z_l \partial z_k} = \frac{\partial^2 W(z)}{\partial z_k \partial z_l}$, yields

$$\frac{g_l'(z_l)}{g_k'(z_k)} = \frac{\partial \mathbb{E}[X_k|z] / \partial z_l}{\partial \mathbb{E}[X_l|z] / \partial z_k}.$$

Every quantity on the right-hand side is observed, so ratios of the derivatives of g are identified.

There are two differences from Allen and Rehbeck (2017) that are worth noting. First, they study a more general model choice structure than ours but focus on the identification of $g(z)$ while treating the distribution of u as a nuisance parameter. We identify this distribution. Second, we shorten their proof substantially by side-stepping the "representative agent's problem" they define.[10]

[9] This assumption substitutes for the non-testable assumption in Allen and Rehbeck (2017) that the second-order cross partials of $\mathbb{E}[\max_{x \in \mathcal{X}} x \cdot (u + g(z)) - t(x)|z]$ are nonzero.

[10] We found this direct proof to be simpler than applying the general conditions of Theorems 1 and 2 in

Theorem 1 and Proposition 3 immediately imply the following result:

Corollary 4. *Suppose Assumptions 1, 2 and 3 and the hypotheses of Proposition 3 are satisfied. If* $v = u + g(z)$ *and* $g(Z)$ *has full support on* $\mathbb{R}^J$*, then* $f_{V|Z}(v|z)$ *is identified.*

5 Application: Combinatorial Auctions

Our empirical application analyzes the combinatorial auction used to procure school lunches in Chile. Section 5.1 describes the data and the institutional setting, and section 5.2 maps the auction into example 4. The analysis then proceeds in two parts. Sections 5.3 through 5.5 estimate win probabilities and payments, bound bidder types, and bound the gains from an efficient allocation, none of which requires an assumption on the distribution of costs. Sections 5.6 and 5.7 then parameterize that distribution and use it to estimate profit margins and to simulate a first-best benchmark.

5.1 Empirical Setting

The National Board for School Aid and Scholarships – henceforth "the auctioneer" – contracts with private catering companies – henceforth "bidders" – to prepare meals and deliver them to schools in the 90 territorial units – henceforth "units". Epstein et al. (2002) and Kim et al. (2014) provide institutional details, and we continue to abbreviate the latter as KOW.

Each year, the auctioneer uses a combinatorial auction to procure meals for about a third of the units. A bid consists of a list of units – henceforth "package" – and a per-meal quote. Bidders qualify by meeting technical and financial requirements, and their qualifications determine which packages they may bid on. KOW documents these requirements in detail. The auctioneer determines the contract allocation by solving an integer linear program that minimizes the total procurement cost and ensures contracts for all auctioned units. Although a bidder may submit many bids, each bidder is awarded at most one of its submitted bids.

We use the same dataset as in KOW. This dataset contains all the bids submitted to the auctions, unit-specific geographic and demographic information and bidder-specific financial and technical ratings. We focus on the 2003 auction to make our results comparable to KOW.

Table 1 describes the auction environment and the outcome. The auctioneer allocated 32 units, totaling about 81 million meals. There were twenty bidders, which together bid on

Allen and Rehbeck (2017). A cost of our approach is that we forego the appealing interpretation of the observed quantity $\mathbb{E}[X|z]$ as the solution to the representative agent's problem.

Table 1: Auction Characteristics

Panel A: Auction environment	
Total meals allocated (millions)	80.76
# units allocated	32
# bidders	20
# bids submitted	43,136
Panel B: Awarded allocation	
Payment (pesos per meal)	410.96
# winning bidders	9

Note: In these tables and elsewhere, the number of meals refers to the total number of meals served per year. Winning bidders are allocated three-year contracts.

Table 2: Units, Bidders, and Bids

	Min	25th percentile	Median	75th percentile	Max
		Panel A: Units (n = 32)			
Meals (millions)	1.64	2.07	2.40	3.09	3.79
# bidders on unit	14	17	18	19	19
# bids on packages containing unit	1,008	4,219	7,204	9,370	11,198
		Panel B: Bidders (n = 20)			
Max meals allowed (millions)	25.92	35.11	37.77	38.77	50.02
Max units allowed	1	3	4	8	8
# units bid on	9	29	32	32	32
# packages bid on	17	505	1,164	3,679	9,157
# feasible packages	32	2,124	10,906	31,351	31,351
		Panel C: Bids (n = 43,136)			
Price (pesos per meal)	299.59	401.92	421.41	440.35	690.65
# units in package	1	4	5	7	8
Package volume (millions of meals)	1.64	9.39	12.82	17.09	25.61
Package density (region HHI)	0.22	0.36	0.50	0.58	1.00

Note: The unit of observation is the unit in Panel A, the bidder in Panel B, and the bid (i.e., the bidder-package pair) in Panel C.

over 43,000 packages. Nine bidders won at least one unit. Across all bidder-package pairs, the average bid was 423 pesos per meal or 71 U.S. cents per meal.[11]

Table 2 describes heterogeneity across units, bidders, and bids. Units vary in the number of package bids containing them, ranging from just over 1,000 to over 11,000, and in their total number of meals, and each unit receives at least one bid from multiple bidders. Bidders face varying limits on the numbers of meals and units they may be allocated, and each bids on far fewer packages than are feasible for it. The bid price per meal exhibits significant heterogeneity across bidder-package pairs. The meal-weighted average price in the awarded allocation, 411 pesos per meal, falls in the 35th percentile of this distribution. The submitted bids also vary in package volume (total meals) and package density (the geographic concentration of units within the package). We define the density of a package as the probability that two meals, drawn independently from the package's distribution of meals across geographic regions, come from the same region.[12] Package volume and density are two potential sources of across-unit cost complementarities, and they enter into the bid and cost specifications below.

5.2 Empirical Model

The Chilean combinatorial auction is a special case of example 4, mechanisms with private information, independent types, and quasilinear utilities. Fitting it into that example requires four objects, namely the action set $\mathcal{A}_i$, the expected outcome function $x_{\mathcal{A}}$, the expected payment function $t_{\mathcal{A}}$, and the payoff type v_i. We construct each in turn.

Mechanism: Each unit, indexed by $k \in \{1, \ldots, K\}$, must be allocated to one of the bidders in the auction, indexed by $i \in \{1, \ldots, I\}$. Let $j \in \{1, \ldots, J\}$ index a package consisting of a subset of units, where $J = 2^K - 1$ is the number of possible nonempty packages. Bidder i's action $a_i = \{a_{ij}\}_j \in \mathcal{A}_i \subseteq (\mathbb{R}_+ \cup \{\infty\})^J$ specifies a bid for each package. The feasible action set $\mathcal{A}_i$ permits finite bids on at most 10,000 packages. We adopt the convention that $a_{ij} = \infty$ if bidder i does not bid on package j.

The combinatorial auction can be represented using an allocation rule $\tilde{x}_i(a_i, a_{-i})$ and a transfer function $\tilde{t}_i(a_i, a_{-i})$. Winning bidders are paid $\tilde{t}_i(a_i, a_{-i}) = -a_i \cdot \tilde{x}_i(a_i, a_{-i})$, the amount of their winning bids. The chosen allocation is determined by finding the lowest total

[11] At the end of 2003, the exchange rate was 599 Chilean pesos per U.S. dollar (OECD, 2016).

[12] The 32 units in the 2003 auction are partitioned into five regions defined by the auctioneer. Our density definition can also be interpreted as a Herfindahl-Hirschman index (HHI) for the package's meal concentration across regions.

price at which all units can be served, which is the solution to the linear integer program

$$\tilde{x}\left(a_1, \ldots, a_I\right) \in \underset{\tilde{x} \in \tilde{\mathcal{X}}}{\operatorname{argmin}} \sum_i \sum_j \tilde{x}_{ij} a_{ij}, \tag{8}$$

where $\tilde{x}_{ij}$ indicates whether bidder i is awarded package j and $\tilde{\mathcal{X}} \subseteq \{0,1\}^{IJ}$ is a set of allowable allocations. The allowable allocations require that every unit be served, that no bidder win more than one package, and that bidder- and region-specific limits on meals and units be respected. All of these constraints are linear in $\tilde{x}_{ij}$, and appendix B.1 states them in full. The solution to the problem is generically unique.

Bidder Types and Equilibrium: Let c_{ij} denote bidder i's cost of supplying package j and let $c_i = \{c_{ij}\}_j$. In the notation of section 2.1, bidder i's valuation vector is $v_i = -c_i$. We assume independent private valuations conditional on observables z_i. In our application, z_i collects observable bidder characteristics, specifically the bidder's size (small or large) and incumbency status.

Assumption 5. *The cost vector of each bidder i, denoted c_i, is independent of the costs of the other bidders and is distributed according to $F_{C|z_i}$. Each bidder's valuation is private information. Bidder characteristics z_i and conditional cost distributions $F_{C|z_i}$ are common knowledge.*

This cost distribution is high-dimensional since $c_i \in \mathbb{R}^J$, where $J = 2^K - 1$ is exponential in the number of units, K. The empirical implementation in section 5.4 will therefore require us to reduce the dimension of the problem.

Before submitting her bid, a bidder observes her own type and characteristics, as well as her competitors' characteristics. We assume that bidders' strategies $\sigma\left(c_i; z_i, z_{-i}\right) : \mathbb{R}^J \to \Delta\mathcal{A}_i$ are described by a type-symmetric Bayesian Nash Equilibrium.

As described in section 3.2, a change of variables in equation (1) yields bidder i's expected outcome function

$$x_{\mathcal{A}}\left(a_i\right) = \int \tilde{x}_i\left(a_i, a_{-i}\right) \mathrm{d}F_{A_{-i}|z_i, z_{-i}}\left(a_{-i}\right) \tag{9}$$

and (pay-as-bid) expected payment function

$$t_{\mathcal{A}}\left(a_i\right) = -a_i \cdot x_{\mathcal{A}}\left(a_i\right). \tag{10}$$

Bidder i's expected outcome is a vector $x_i = \{x_{ij}\}_j \in \mathcal{X}_i \subseteq [0,1]^J$ specifying its probability of being allocated each package j. The function (9) integrates over the distribution of bidder

i's opponents' bids conditional on observables, $F_{A_{-i}|z_i,z_{-i}}(a_{-i}) = \prod_{i' \neq i} F_{A_{i'}|z_i,z_{-i}}(a_{i'})$, where the equality follows from the independence of costs in assumption 5. Equilibrium implies that for each bidder i, the bid vector a_i satisfies

$$a_i \in \underset{a_i' \in \mathcal{A}_i}{\operatorname{argmax}} (a_i' - c_i) \cdot x_{\mathcal{A}}(a_i') \tag{11}$$

5.3 Win Probabilities $x_{\mathcal{A}}$ and Payments $t_{\mathcal{A}}$

Computing the expected outcome function $x_{\mathcal{A}}(a)$ and expected payment function $t_{\mathcal{A}}(a)$ requires estimates of the distribution of opponent bids, $F_{A_{-i}|z_i,z_{-i}}(a_{-i}) = \prod_{i' \neq i} F_{A_{i'}|z_i,z_{-i}}(a_{i'})$. We describe our approach in this subsection.

5.3.1 Estimation Procedure

Without further restrictions, the bid vectors A_i are high-dimensional because the number of components equals the number of packages. We follow KOW when modeling and estimating the distribution $F_{A_i|z_i,z_{-i}}(a_i)$. The main restrictions are that (i) each bidder knows the (exogenous) subset of packages on which its competitors will bid, and (ii) beliefs about the bid distribution follow a low-dimensional parametric distribution.

Our parametric specification is motivated by the economics of combinatorial auctions. These auctions attempt to exploit complementarities in costs. Package bidding lets a bidder avoid winning only part of a package that it values as a whole, and how much it helps depends on how large the complementarities are, how much they vary across bidders, and how far they pass through into bids. Thus, the model allows submitted bids to vary systematically with package volume and geographic density. These bid adjustments may reflect underlying cost complementarities or substitutabilities, variation in markups, or both.

Bidder i's per-meal bid on package j (from the perspective of i's competitors) is specified as a volume-weighted sum of unit-specific **base prices** plus **bid adjustments** that capture scale and scope discounts due to package volume and density:

$$\frac{a_{ij}}{q_j} = \sum_{l=1}^{L^{\text{volume}}} \mathbf{1}\{q_j \in \mathcal{Q}_l\} \beta^{\text{volume}}_{l,\text{size}_i} + \sum_{l=1}^{L^{\text{density}}} \mathbf{1}\{d_j \in \mathcal{D}_l\} \beta^{\text{density}}_{l,\text{size}_i} + \sum_{k \in j} \frac{q_k}{q_j} \tilde{\beta}^{\text{unit}}_{ik} + \epsilon_{ij}, \tag{12}$$

where $q_j = \sum_{k \in j} q_k$ is the package's total volume (in millions of meals per year) and d_j is

our measure of the geographic concentration of units within the package.[13] The unit-specific base prices are distributed as

$$\tilde{\beta}_{ik}^{\text{unit}} = \beta_k^{\text{unit}} + \beta^{\text{incumb}} \text{incumb}_{ik} + \omega_{ik}, \tag{13}$$

where incumb_{ik} is an indicator for whether bidder i is the incumbent supplier for unit k and $\omega_i = (\omega_{i1}, \ldots, \omega_{iK}) \sim N(0, \Sigma_\omega)$ allows for correlation across units within a bidder. Volume and density are discretized into bins, and the bid adjustment amounts depend on bidder size.[14]

Our empirical parameterization uses $L^{\text{volume}} = 4$ volume economies, $L^{\text{density}} = 3$ density economies, and 32 unit-specific costs. Volume bins are cut at 6, 12, and 18 million meals per year and density bins at HHI values of 0.5 and 1, with the lowest volume bin and the highest density bin omitted.[15]

It is convenient to write this specification in matrix form.

Definition 3. Let M be the $J \times L$ matrix whose row for package j has entry q_j in the columns for the volume and density bins that contain j and entry q_k in the column for each unit $k \in j$.

Equation (12) then reads $a_i = M\beta_i + \varepsilon_i$, where β_i stacks the bid adjustments and the unit base prices that apply to bidder i and ε_i collects the package-level shocks. The same matrix M parameterizes costs in section 5.4.

We estimate the parameters θ_a of the bid distribution in two stages by feasible generalized least squares, first from equation (12) and then from equation (13), taking the first-stage base price estimates as data.

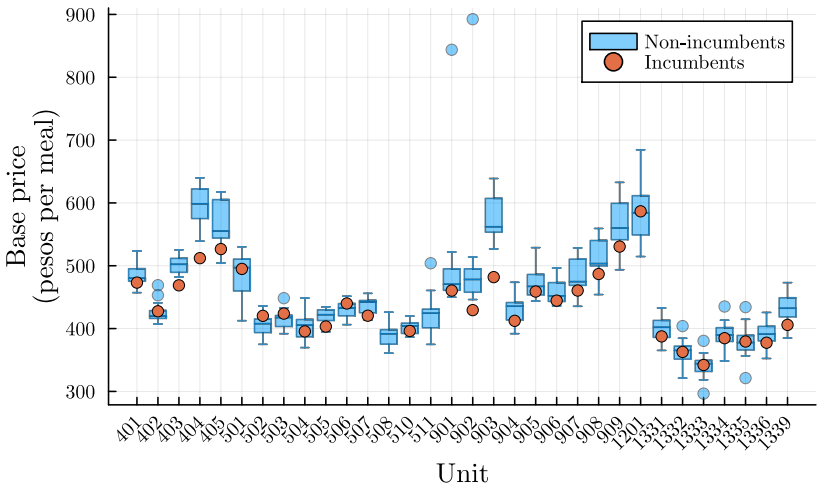
Given bid distribution parameter estimates $\hat{\theta}_a$, we estimate the outcome function $x_{\mathcal{A}}(\cdot)$ and payment function $t_{\mathcal{A}}(\cdot)$ via simulation of equations (9) and (10), respectively. For each bidder i, we use $S_a = 1{,}000$ competitor bids a_{-i} from the distribution $F_{A_{-i}|z_i,z_{-i}}\left(\cdot; \hat{\theta}_a\right)$. This simulation step differs from the empirical approach to multiunit auctions in Kastl (2011) and

[13] Specifically, the density of package j is $d_j = \sum_r \left(\frac{\sum_{k \in j \cap r} q_k}{q_j} \right)^2$, where the units have been partitioned into regions indexed by $r \in \{1, \ldots, R\}$.

[14] We call a bidder small if it is allowed to win at most six units and large if it is allowed to win more than six. Our bid distribution parameterization largely parallels that of KOW. We define package density differently and allow the variance-covariance matrix Σ_ω to be unrestricted.

[15] The volume and density bins are right-open intervals. The omitted highest density bin contains packages whose meals all fall within a single region.

Figure 2: Distribution of Bidders' Unit Base Prices



Note: The blue boxplots show the distribution of nonincumbent bidders' base prices charged for each unit, denoted β_k^{unit} in equation (12). Boxes show the 25th percentile, median, and 75th percentile. Lower whiskers extend to the lowest observed data point that is within a distance of 1.5 times the interquartile range (IQR) from the 25th percentile. Likewise, upper whiskers extend to the highest observed data point within 1.5 times the IQR from the 75th percentile. Blue dots indicate outliers. Each orange dot indicates the base price bid by the incumbent bidder serving that unit.

Cassola et al. (2013): instead of resampling from competitors' bids, we draw competitor bid profiles a_{-i} from the estimated distribution $F_{A_{-i}|z_i,z_{-i}}$.

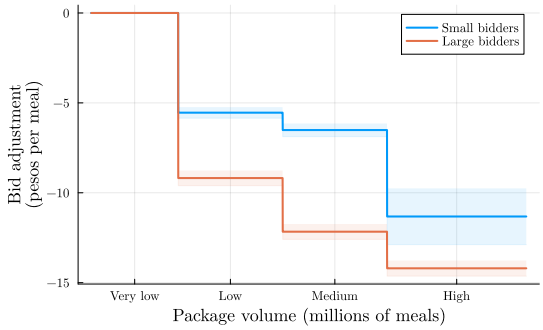
5.3.2 Estimates

Bid Distribution: Figure 2 plots the distribution of unit-specific base prices. Across all bidders and all units, the mean unit-specific base price is 450 pesos per meal, though this masks considerable variation, particularly across units. The mean base price for unit k, denoted β_k^{unit} in equation (13), ranges from 342 to 592 pesos per meal. Within a unit, bidders' base prices are correlated. However, we estimate that the incumbent bidder in a unit submits bids for that unit that are 19.00 pesos per meal lower on average than those of its nonincumbent competitors, suggesting a potential incumbency cost advantage.

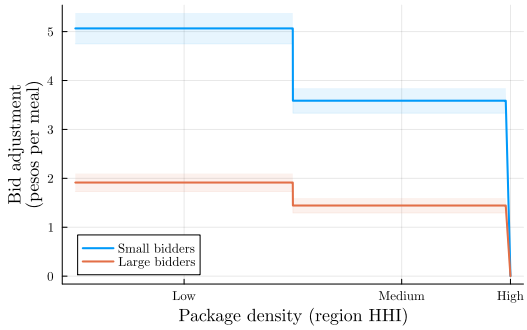
Figure 3 shows the estimated bid adjustment functions for small and large bidders. We find

Figure 3: Estimated Bid Adjustment Functions

(a) Volume



(b) Density



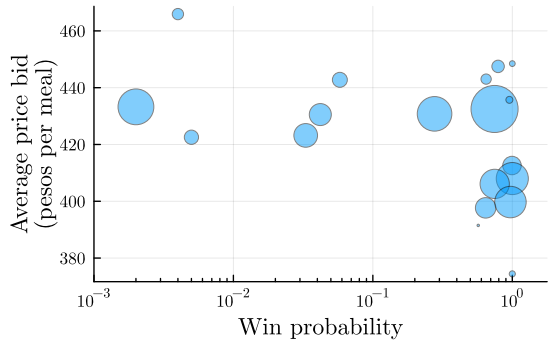
Note: Figures show estimated bid adjustments due to economies of volume and density. Shaded areas show 95 percent confidence intervals computed using the standard errors of the estimated coefficients. The lowest volume bin and the highest density bin are omitted base categories, so bid adjustments are in comparison to low-volume, high-density packages – for example, single-unit packages.

evidence of significant downward bid adjustments due to volume and density. Bid adjustments become more negative as package volume and density increase. We estimate that the prices bid on the largest (highest-volume) packages we observe are, on average, 11.3 to 14.2 pesos per meal lower than those on the smallest packages. Likewise, we estimate that average bids on low-density packages are 1.9 to 5.1 pesos per meal higher than those on high-density ones. Small bidders offer smaller discounts for package volume than large bidders do, though the pattern reverses for density, where small bidders offer the greater discounts. Figure B.1 compares the estimated bid distribution to the observed one, and the fit is good.

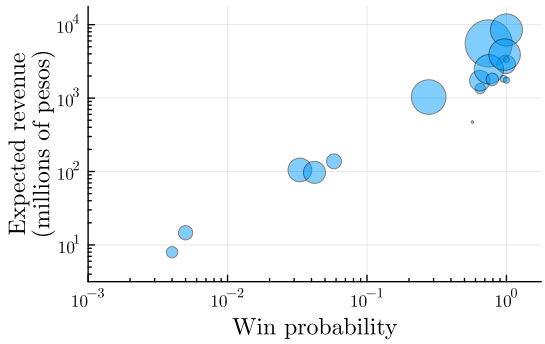
Outcome and Payment Functions: Figure 4 plots bidders' win probabilities against their average bids and their expected revenues. Bidders vary substantially in their probabilities of winning at least one unit. Two bidders are estimated to win with zero probability at their submitted bids, while the win probabilities of another four bidders are estimated to be greater than 95 percent. The left panel shows that bidders who submit lower average bids tend to have a greater probability of winning a package, and the right panel shows that expected revenues $-t_{\mathcal{A}}(a_i)$ increase in win probabilities. Win probabilities are, perhaps surprisingly, not very correlated with the *number* of bids submitted.

Figure 4: Bidders' Average Prices, Win Probabilities, and Expected Revenues

(a) Win Probability vs. Average Price Bid



(b) Win Probability vs. Expected Revenue



Note: Each circle represents a bidder. The marker size is proportional to the number of bids submitted by that bidder. The horizontal axis indicates the bidder's probability of winning any package – that is, the sum of the elements of $x_a(a_i)$, the expected allocation vector. In the left panel, the vertical axis indicates the bidder's average per-meal price bid. In the right panel, the vertical axis indicates the bidder's expected revenue, $t_a(a_i)$. These figures omit two bidders which are estimated to have zero win probability.

5.4 Estimated Bounds on Types

This section applies the identification results from section 4 to bound bidder types nonparametrically. These bounds are an instance of partial identification via random sets: they are the capacity inequalities of the optimal-choice correspondence, and the outer rationalizing set we report is an outer region in the sense of Beresteanu et al. (2011). We first describe how to reduce the dimension of the type space and compute bounds on each bidder's type from observed bids.

5.4.1 Restricting the Dimension of the Type Space

Bounding types requires a cost for every package a bidder could have won. The most flexible specification would leave each c_{ij} unrestricted, permitting arbitrary complementarities and substitutabilities among a package's constituent units. That is not feasible here. The number of packages is $J = 2^K - 1$, which grows exponentially in the number of units, and with $K = 32$ it exceeds 4.2 billion. Computing the revealed preference bounds $\partial t\,(a_i)$ at that scale is out of reach, and estimating $F_{C|z_i}$ would be harder still.

We therefore give costs the low-dimensional structure that the bid specification in equation

(12) imposes on bids. Specifically, we assume that bidder i's cost vector is

$$c_i = M\gamma_i, \qquad \gamma_i = \left(\left(\gamma_{il}^{\text{volume}}\right)_l, \left(\gamma_{il}^{\text{density}}\right)_l, \left(\gamma_{ik}^{\text{unit}}\right)_k\right) \in \mathbb{R}^L, \tag{14}$$

with M as in definition 3 and $L \ll J$. Here $\gamma_{il}^{\text{volume}}$ is the per-meal cost adjustment that applies when the package falls in volume bin l, $\gamma_{il}^{\text{density}}$ is the adjustment when it falls in density bin l, and $\gamma_{ik}^{\text{unit}}$ is bidder i's per-meal cost of supplying unit k. The first two capture economies or diseconomies of scale and of density.

This approach contrasts with that in KOW, which solves the dimensionality problem by assuming that the markups that are bid, which are endogenous, are a low-dimensional function of package characteristics. Rather than restricting endogenous markups, we place restrictions directly on the distribution of bidders' costs. These restrictions will allow us to ease the computation of types consistent with optimal bidding.

5.4.2 Bounding Types

The optimality condition (11) substituted with $c_i = M\gamma_i$ forms the basis for our estimation strategy. Given that bidder i submits a_i, the revealed preference inequality (15) implies that for each deviation $a_i' \in \mathcal{A}_i$,

$$M\gamma_i \cdot \left[x_{\mathcal{A}}\left(a_i\right) - x_{\mathcal{A}}\left(a_i'\right)\right] \leq a_i \cdot x_{\mathcal{A}}\left(a_i\right) - a_i' \cdot x_{\mathcal{A}}\left(a_i'\right). \tag{15}$$

Thus, for each deviation a_i', this inequality specifies a half-space in the domain of types. The revealed preference bounds $\partial t\left(a_i\right)$, the intersection of these half-spaces for all deviations in $\mathcal{A}_i$, forms a convex polyhedral set.[16]

We denote our estimate of the set of rationalizable types for bidder i by $\mathcal{G}_i$. Two implementation challenges shape how we compute it. First, evaluating a single deviation requires simulating $x_{\mathcal{A}}\left(\cdot\right)$ at the alternative bid, which takes $S_a = 1{,}000$ solutions of a high-dimensional integer program, so checking every deviation is infeasible.[17] Second, $x_{\mathcal{A}}\left(\cdot\right)$ is estimated with

[16] Our model with this simplified type space can be recast in terms of our original model. Observe that $v_i \cdot x = -M\gamma_i \cdot x = -\gamma_i \cdot \left(M^T x\right)$. Because x is a vector of win probabilities, $M^T x$ is a vector with the expected columns of M under assignment probabilities x. Let $\hat{x} = M^T x$; thus, $v_i \cdot x = -\gamma_i \cdot \hat{x}$. Define $t_{\hat{X}}\left(\hat{x}\right) = \inf\left\{t_{\mathcal{A}}\left(a\right),\ a \in \mathcal{A} : \left(M^T x_{\mathcal{A}}\left(a\right)\right) = \hat{x}\right\}$ where $t_{\mathcal{A}}\left(a_i\right) = -a_i \cdot x_{\mathcal{A}}\left(a_i\right)$. The bidder's problem therefore maximizes $-\gamma_i \cdot \hat{x} - t_{\hat{X}}\left(\hat{x}\right)$.

[17] Checking many deviations also results in a complex description of the set $\mathcal{G}_i$. As the number of half-spaces grows, polyhedral operations like checking set membership, projection onto subspaces, and elimination of dimensions become more difficult.

simulation error. Larger deviations are estimated more precisely but yield a larger than necessary $\mathcal{G}_i$ (Fu, 2006).

We compute $\mathcal{G}_i$ from a subset of deviations $\tilde{\mathcal{A}}_i \subset \mathcal{A}_i$. Each deviation changes bidder i's quote on a single package j drawn from a set $\tilde{\mathcal{J}}_i$ that spans the type dimensions and includes all single-unit packages, holding its quotes on other packages fixed. A downward deviation lowers the quote to the level that would newly win j in 5 percent of the simulation draws in which i lost j at its observed bid, and an upward deviation raises it to the level that would newly lose j in 50 percent of the draws in which i won it. We also impose three constraints on $\mathcal{G}_i$, namely (i) costs lie below bids ($M\gamma_i \leq a_i$), (ii) the cost of each package on which bidder i bids is nonnegative ($M\gamma_i \geq 0$), and (iii) the cost of each single-unit package is nonnegative regardless of whether bidder i bid on it ($\gamma_i^{\text{unit}} \geq 0$). Appendix B.3 describes the construction of $\tilde{\mathcal{J}}_i$ and the computation of the threshold quotes.

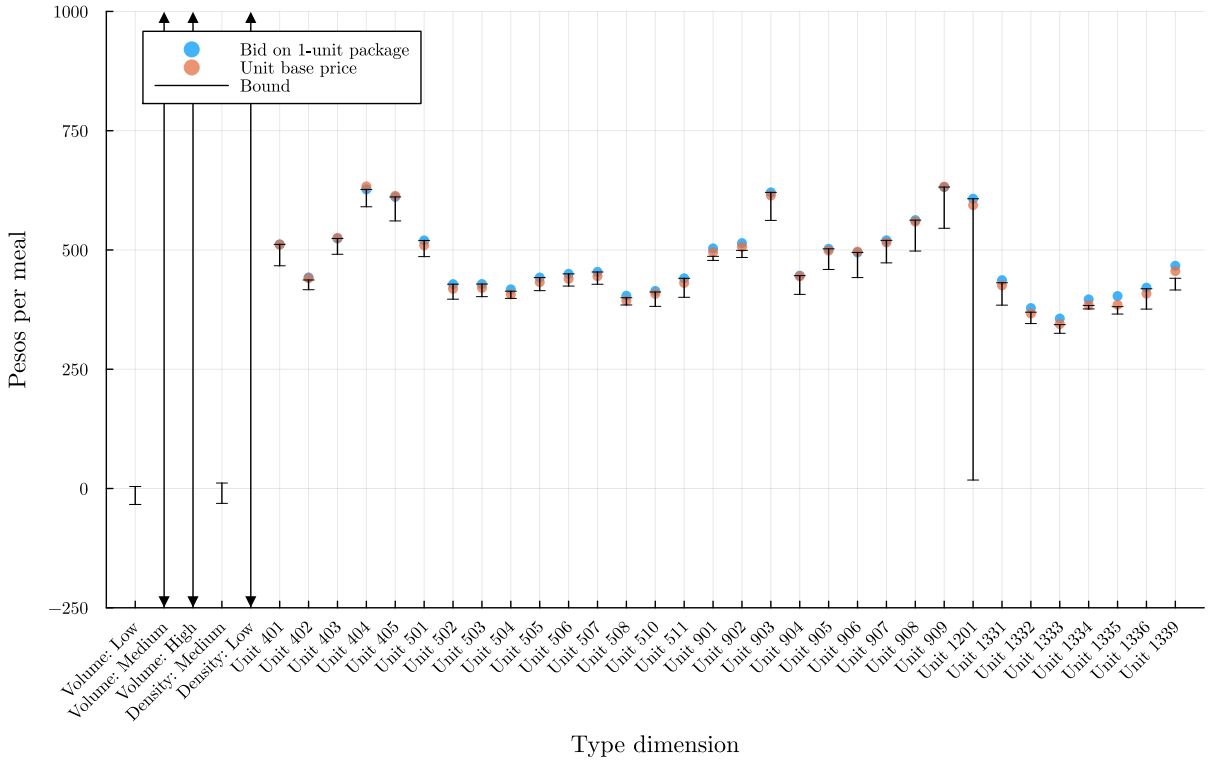
The revealed preference inequalities and the three additional constraints yield a convex polyhedral set. Each downward deviation implies a lower bound on some linear combination of type dimensions, while each upward deviation provides an upper bound. Constraint (i) generates additional upper bounds, and constraint (ii) generates additional lower bounds. Single-unit packages in the set $\tilde{\mathcal{J}}_i$ and constraint (iii) ensure that unit-specific costs are bounded even if bidder i does not bid on certain units. Because we evaluate only a subset of the feasible deviations, $\mathcal{G}_i$ is an outer set containing the revealed preference bounds, $\partial t\left(a_i\right) \subseteq \mathcal{G}_i$. Working with a superset is conservative, since it widens the recovered bounds rather than risking exclusion of the true type.

5.4.3 Estimates

Figure 5 shows type bound estimates for one bidder. Each black whisker in figure 5 plots a projection of the polyhedral set $\mathcal{G}_i$ onto a single dimension. This example bidder is not allowed to win any packages involving medium and high volume bins or the low density bin, so we are unable to bound its realizations of those dimensions. The estimated bounds suggest that this bidder has economies of volume, though we cannot rule out very small diseconomies of volume.[18] Likewise, this bidder exhibits either small economies or small diseconomies of density. Because single-unit packages incur no volume or density economies, the bidder's bids on such packages (the blue dots in figure 5) bound the corresponding unit-specific costs

[18] Higher values for the type dimensions $\gamma_{il}^{\text{volume}}$ and $\gamma_{il}^{\text{density}}$ increase the per-meal package cost in equation (14). Hence, a more negative value for $\gamma_{il}^{\text{volume}}$ indicates a larger economy of volume, while a more positive value indicates a larger *diseconomy* of volume.

Figure 5: Example Bidder Type Estimates



Note: This figure plots type estimates for an example bidder i. For each of the unit-specific costs $\gamma_{ik}^{\text{unit}}$, the blue dot indicates bidder i's bid on the single-unit package containing only unit k. The orange dot indicates bidder i's base bid on unit k, denoted $\tilde{\beta}_{ik}$. Each black whisker is the projection of the distribution-free type bounds $\mathcal{G}_i$ onto a single dimension. Figure B.3 in the appendix overlays the estimated type distribution on the same bounds.

from above. These single-unit package bids are very similar but not identical to the unit k base prices $\tilde{\beta}_{ik}^{\text{unit}}$ from section 5.3.1 (in orange in figure 5), which are estimated using data from all of i's package bids containing unit k.

5.5 Bounds on Efficiency Gains

This section bounds the cost inefficiency of the observed combinatorial auction allocation using only the bounds on bidder types described in the previous section. The output is a pair of frontiers that trace the smallest and the largest efficiency gains consistent with the bounds. In the case of a combinatorial auction problem, the efficient allocation can be

obtained in principle using the strategy-proof Vickrey-Clarke-Groves (VCG) mechanism with the pivot payment rule.

We compare the observed allocation to the efficient allocation, which minimizes the total cost of supplying all units. The efficient allocation solves

$$\tilde{x}\left(c_1, \ldots, c_I\right) \in \underset{\tilde{x} \in \tilde{\mathcal{X}}}{\operatorname{argmin}} \sum_i \sum_j \tilde{x}_{ij} c_{ij} \tag{16}$$

where $\tilde{\mathcal{X}}$ is the same set of allowable allocations imposed in the combinatorial auction. The combinatorial auction (8) minimizes total payments, which are the sum of winning bids, while (16) minimizes the total cost of supply. Both are subject to the allocation constraints imposed in Chile. We restrict the set of bidder-package pairs (i, j) in (16) to those in which we observe i submitting a bid on j.[19,20]

Since $c_i = M\gamma_i$, the vector of costs of all bidder-package pairs is determined by $\gamma = \left(\gamma_1^\top, \ldots, \gamma_I^\top\right)^\top \in \mathbb{R}^{IL}$, which stacks all the bidder types. Write the gain from reallocating conditional on γ as

$$\Delta C\left(\gamma\right) = \max_{\tilde{x} \in \tilde{\mathcal{X}}} \sum_i \sum_j \left(M\gamma_i\right)_j \left(\tilde{x}_{ij}^{OA} - \tilde{x}_{ij}\right), \tag{17}$$

where $\tilde{x}^{OA}$ denotes the observed allocation. This is the amount by which the cost of supplying all units would fall if types were γ and packages were efficiently assigned. The bounds $\mathcal{G} = \prod_i \mathcal{G}_i$ then imply lower and upper bounds on the efficiency gains:

$$\underline{\Delta C} = \min_{\gamma \in \mathcal{G}} \Delta C\left(\gamma\right), \tag{18}$$

$$\overline{\Delta C} = \max_{\gamma \in \mathcal{G}} \Delta C\left(\gamma\right). \tag{19}$$

The two bounds are hard to compute for different reasons. The objective function $\Delta C(\gamma)$ is

[19] We make this restriction because computing the efficient allocation requires solving a high-dimensional integer program. With $I = 18$ bidders (excluding the two low-quality bidders discussed in footnote 20) and $K = 32$ units, there are $I \times \left(2^K - 1\right)$, or over 77 billion, bidder-package pairs. Even after dropping infeasible bidder-package pairs, we are still left with over 100 million decision variables. In contrast, there are only 43,136 pairs in which we observe i submitting a bid on j.

One interpretation of this restriction is that a bidder submits a bid on every package it is seriously prepared to supply, in which case it is without loss. Another is that it is an approximation. Either way, the restriction weakly increases the computed cost of the efficient allocation and therefore weakly understates the efficiency gains.

[20] Following KOW, we also exclude two bidders who submit exceptionally low bids. KOW states that "Despite their competitive prices, these bidders did not win any units [in the observed auction] and were disqualified from the allocation process because of quality considerations." We include these two low-quality bidders in the estimation of the bid and type distributions because their competitors likely did not know they would ultimately be disqualified.

convex because it is a maximum of functions that are linear in γ. Although the lower bound problem in equation (18) minimizes a convex function over a convex set, it is a saddle point problem in which evaluating the objective at a single γ requires solving an integer program over $\tilde{\mathcal{X}}$. Thus, the usual route of dualizing that inner maximization is unavailable because relaxing the integrality constraints in $\tilde{\mathcal{X}}$ does not recover its convex hull. The upper bound problem in equation (19) maximizes a convex function over a convex set. Its maximum is therefore attained at an extreme point of $\mathcal{G}$, but $\mathcal{G}$ has many extreme points.

Appendix B.6.1 details how we solve the two problems. Writing $\Delta C(\gamma)$ through its epigraph instead turns the lower bound into a linear program with one constraint per feasible allocation, which we solve exactly by generating constraints as they bind. For the upper bound, we search over extreme points of $\mathcal{G}$ and bound $\overline{\Delta C}$ tightly from below and above.

Each type profile in $\mathcal{G}$ implies costs for both the observed allocation and the efficient allocation. The bounds $\mathcal{G}$ therefore trace out a set of cost pairs that the bids cannot rule out. Figure 6 plots the boundary of that set together with the 45-degree line. The potential efficiency gain at any point is its vertical distance below the 45-degree line. Given $\mathcal{G}$, the cost of the observed allocation itself lies between 240.0 and 410.0 pesos per meal. For each value between these two extremes, the minimum-gains frontier reports the smallest efficiency gain consistent with the bounds and the maximum-gains frontier the largest. We obtain these frontiers by solving the programs in equations (18) and (19) with the additional constraint that the per-meal cost of the observed allocation equals a given value $\bar{c} \in [240.0, 410.0]$, that is, $\sum_i \sum_j (M\gamma_i)_j \tilde{x}_{ij}^{OA} = \bar{c} \sum_k q_k$.

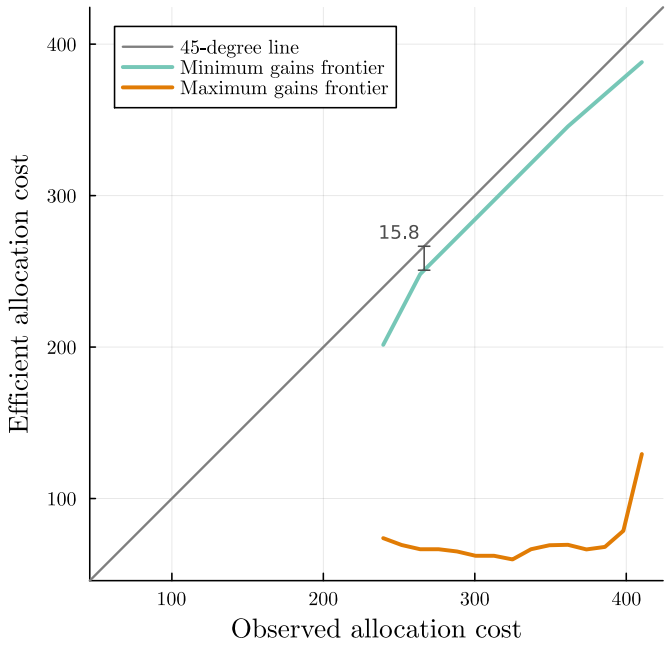
The minimum inefficiency of the observed combinatorial auction allocation across all types $\gamma \in \mathcal{G}$ is 15.8 pesos per meal, or 3.8 percent of the total payment to winning bidders. The maximum-gains frontier in figure 6 traces the largest gains consistent with the estimated bounds. Because those bounds are wide, it primarily reflects what the data leave unidentified rather than an economic estimate of the gains.

Our results therefore show that production costs would fall by at least 3.8 percent of the total payment to winning bidders if the VCG pivot mechanism could be used, although there are important practical barriers to its implementation (Ausubel and Milgrom, 2006).

5.6 Estimated Type Distribution

The type bounds in section 5.4 are nonparametric. This section shows how one can obtain point estimates of the distribution of cost types $F_{\Gamma_i|z_i}$ under additional parametric assump-

Figure 6: Bounds on Observed and Efficient Allocation Costs



Note: The minimum- and maximum-gains frontiers, which are the distribution-free bounds on the cost of the observed and efficient allocations, are shown together with the 45-degree line.

tions.

We parameterize the type distribution $F_{\Gamma_i|z_i}$ as a function of bidder size and incumbency. Conditional on these observables, we assume that the type distribution is multivariate normal with unrestricted variance-covariance matrix Σ_ν. The mean values of the volume and density components are $\mu^{\text{volume}}_{\text{size}_i}$ and $\mu^{\text{density}}_{\text{size}_i}$ respectively, again potentially differing by bidder size. Mean unit-specific costs are given by $\mu^{\text{unit}} + \mu^{\text{incumb}}\text{incumb}_i$, the sum of a common mean vector and an incumbency shifter. We estimate the parameter vector $\theta_\gamma = \left(\mu^{\text{volume}}, \mu^{\text{density}}, (\mu^{\text{unit}}_k)_k, \mu^{\text{incumb}}, \Sigma_\nu\right)$ using a Gibbs sampler given bidder i's type bounds $\mathcal{G}_i$.[21] The likelihood is given by[22]

$$L\left(\theta_\gamma \mid \mathcal{G}_i\right) = \int 1\left\{\gamma_i \in \mathcal{G}_i\right\} \mathrm{d}F_{\Gamma_i|z_i}\left(\gamma_i; \theta_\gamma\right). \tag{20}$$

Appendix B.4.1 describes the Gibbs sampler in greater detail.

Table 3 reports posterior means and 95 percent credible intervals for the volume and density economy parameters.[23] The mean economies μ are imprecisely estimated, and every credible interval for a mean contains zero. Their signs nonetheless line up with the bid adjustments in section 5.3.2. Mean volume economies are negative for large bidders in every bin, and the mean density parameters are positive throughout, so packages spread across regions cost more to supply than concentrated ones.

[21] This algorithm generates a Markov chain which converges to the posterior distribution. We take the posterior mean parameters as our point estimate $\hat{\theta}_\gamma$. We use the posterior mean rather than the mode because the former is more robustly estimated when the parameter space is high-dimensional and the posterior distribution has potentially many local maxima.

[22] Because $\mathcal{G}_i \supseteq \partial t(x_i)$, this is a coarsened likelihood in the sense of Heitjan and Rubin (1991): it conditions on $\gamma_i \in \mathcal{G}_i$ rather than the finer event $\gamma_i \in \partial t(x_i)$. Since $\mathcal{G}_i$ is a larger set, $L(\theta_\gamma \mid \mathcal{G}_i) \geq L(\theta_\gamma \mid \partial t(x_i))$, making the likelihood less informative about θ_γ. Treating $L(\theta_\gamma \mid \mathcal{G}_i)$ as the likelihood requires that the coarsening be at random, i.e., that the selection of γ_i within $\mathcal{G}_i$ is ignorable. We maintain the CAR condition that, conditional on the observed auction environment, the mechanism generating the retained set $\mathcal{G}_i$ does not vary with γ_i within $\mathcal{G}_i$. We impose this assumption here, and nowhere else in the paper, to obtain point estimates.

[23] Table B.4 reports the corresponding estimates for the unit cost and incumbency shifter parameters, and figure B.2 compares the distributions of bidders' unit-specific base bids $\tilde{\beta}_{ik}$ and unit costs $\gamma^{\text{unit}}_{ik}$.

Table 3: Gibbs Posterior Means of Volume and Density Economy Parameters

	μ						σ		
	Small bidders			Large bidders					
		95% credible interval			95% credible interval			95% credible interval	
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
Low volume	-3.64	-18.54	11.46	-9.55	-27.48	8.53	27.37	19.68	37.69
Medium volume	-8.91	-33.32	15.57	-14.68	-38.12	8.23	31.21	21.04	45.88
High volume	9.60	-49.55	75.58	-17.50	-51.40	12.57	38.79	23.67	63.15
Low density	2.68	-22.42	27.75	1.43	-22.25	24.79	33.44	22.50	47.92
Medium density	3.96	-11.64	20.03	6.20	-11.43	23.80	28.06	19.69	39.32

Note: This table presents estimates of $\left(\mu^{\text{volume}}, \mu^{\text{density}}\right)$ and the corresponding diagonal elements of Σ_ν. The 95 percent credible interval (LB, UB) columns report the 2.5th and 97.5th percentiles of the marginal posterior distribution of each parameter.

The dispersion parameters σ are estimated more precisely. They are several times larger than most of the mean estimates and their credible intervals are bounded away from zero, so bidders differ substantially in their economies even where the average economy is not pinned down. This dispersion, rather than any single mean, can be exploited by reallocating packages efficiently.

Estimates from the parametric model, which pool information from multiple bidders, imply that certain cost types satisfying the nonparametric bounds are much more likely than others. For instance, volume economies of up to 33 pesos per meal cannot be ruled out for the example bidder in figure 5. However, figure B.3 in the appendix, which overlays the marginals of the estimated distribution $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$ on the bounds of figure 5 for the same example bidder, shows that the estimated distribution places little mass there and little mass on unit costs above the bidder's own per-meal prices.

The parametric model also implies tighter estimates of economic quantities and counterfactuals of interest. For instance, the type bounds $\mathcal{G}_i$ alone imply that the aggregate profit margin on awarded bids is between 0.2 and 41.7 percent of the total payment to winning bidders. The estimated type distribution conditional on the type bounds implies an expected aggregate margin of 7.8 percent, which is much lower than the upper bound but higher than the KOW estimate of 4.8 percent on awarded bids.[24] Appendix B.5 provides further details on the estimates of the profit margins. The next section will show that the efficiency gains are also estimated quite precisely under the parametric assumptions.

5.7 The VCG Mechanism and Sources of Efficiency Gains

We now simulate counterfactual outcomes under the VCG pivot mechanism, drawing bidder costs from the estimated type distribution conditional on the type bounds. The mechanism implements the efficient allocation of section 5.5 and pays each winning bidder the amount of the positive externality generated by its participation:

$$\tilde{t}_i\left(c_i, c_{-i}\right) = \left(\min_{x \in \tilde{\mathcal{X}}_{-i}} \sum_{i' \neq i} \sum_j x_{i'j} c_{i'j}\right) - \left(\sum_{i' \neq i} \sum_j \tilde{x}_{i'j}\left(c_i, c_{-i}\right) c_{i'j}\right),$$

[24] A potential reason for the difference is that KOW directly parameterizes markups instead of costs. This approach reduces the heterogeneity in markups relative to our model. Thus, some of the difference between our estimates and those in KOW may arise from the decision to estimate markups rather than directly targeting the distribution of costs.

where $\tilde{\mathcal{X}}_{-i} = \left\{x \in \tilde{\mathcal{X}} : x_{ij} = 0 \text{ for all } j\right\}$ is the set of allowable allocations that do not award any package to bidder i. This is the difference between the minimum allocation cost achievable in bidder i's absence and the cost borne by the other bidders in the efficient allocation when i participates.

Table 4 presents simulated allocation costs, producer surpluses, and payments to bidders under the observed and counterfactual VCG allocations. The first column ("Min-max") presents the distribution-free bound as calculated in section 5.5, whereas the other columns report summary statistics from the parametric estimates. The table therefore compares results under the parametric assumptions with those from the revealed preference bounds.

The parametric specification finds economically larger efficiency gains in moving from the observed allocation to the efficient allocation than the nonparametric lower bound. We estimate the efficient allocation cost to be 8.6 percent lower than the observed allocation cost, a gain of 32.6 pesos per meal compared with the lower bound of 15.8 pesos per meal from section 5.5.

The VCG mechanism also results in lower total payments and lower average producer surplus. The producer surplus falls by less than the payments because of the lower cost of supplying the packages.

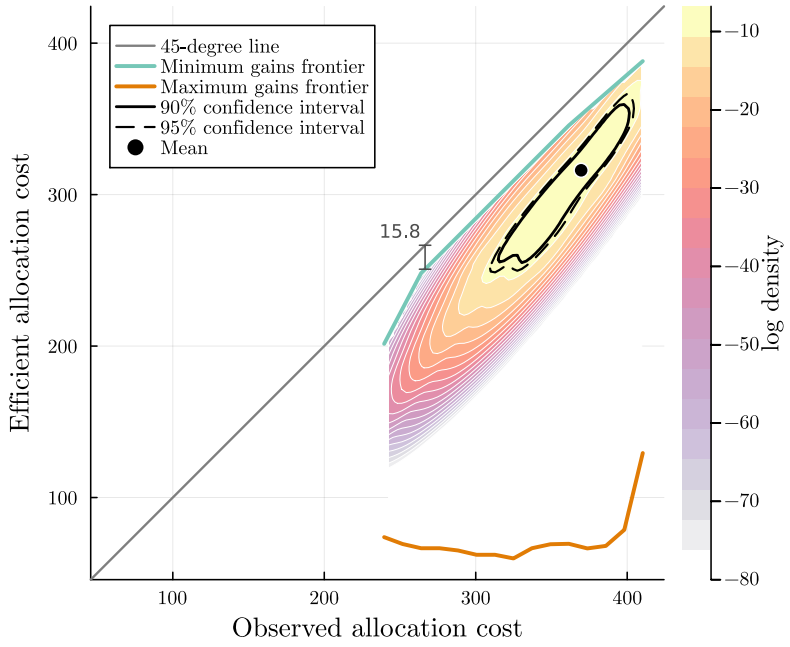
Figure 7 overlays an estimate of the joint density of the observed and efficient allocation costs implied by the estimated type distribution on top of the bounds presented previously in section 5.5. Each draw of the bidders' cost types yields a per-meal cost for the observed allocation and for the efficient allocation, and the contours are those of a kernel-density estimate of these cost pairs, darker where pairs are more likely. Appendix B.6.2 describes the estimator.

We investigate the source of the efficiency gains by recomputing allocation costs for each auction design under alternative assumptions on cost economies, holding fixed the simulated efficient allocations. Table 5 reports the results. In each panel, the first row shows the same baseline allocation cost as in table 4. The second row shuts down heterogeneity in economies of volume and density, and the third row sets these economies to zero.

Bidder-specific heterogeneity in economies of volume and density contributes to the estimated efficiency gains. Replacing each bidder's volume and density cost parameters with their across-bidder and across-draw averages reduces the efficiency gain over the observed allocation from 8.6 percent to 6.6 percent. Setting these cost parameters to zero increases allocation costs under both mechanisms and further reduces the efficiency gain to 6.4 percent.

The results suggest that package bidding in a combinatorial auction is not able to fully realize

Figure 7: Joint Distribution of Observed and Efficient Allocation Costs



Note: Contours show the estimated log-density of the joint distribution of observed and efficient allocation costs implied by the estimated type distribution. The minimum- and maximum-gains frontiers (the distribution-free bounds) and the 45-degree line are overlaid.

Table 4: Observed Allocation vs. VCG Mechanism

	Min-max	Mean	2.5th percentile	97.5th percentile
		Panel A: Observed Allocation		
Cost	266.6	378.8	356.7	393.3
Producer surplus	144.4	32.2	17.7	54.2
Payment	411.0	411.0	-	-
		Panel B: VCG Allocation		
Cost	250.7	346.2	320.2	366.6
Producer surplus	129.6	28.1	2.2	58.3
Payment	380.3	374.3	360.3	386.4
		Panel C: Difference (OA - VCG)		
Cost	15.8	32.6	16.8	44.9
Producer surplus	14.8	4.1	-14.2	25.8
Payment	30.6	36.7	24.6	50.7

Note: Summary statistics of simulated observed allocation and VCG outcomes under the estimated type distributions $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$, conditioning on the cost bounds $\mathcal{G}_i$. All values are in pesos per meal. The Min-max column reports the results from the allocation that minimizes efficiency gains within the identified distribution-free bounds; the remaining columns are computed from the estimated type distribution.

Table 5: Observed and Efficient Allocation Costs Under Alternative Cost Economies

	Min-max	Mean	2.5th percentile	97.5th percentile
	Panel A: Observed Allocation			
Baseline	266.6	378.8	356.7	393.3
Common economies	273.3	391.6	375.0	403.3
Zero economies	268.2	395.0	378.3	406.6
	Panel B: Efficient Allocation			
Baseline	250.7	346.2	320.2	366.6
Common economies	260.2	366.0	348.5	381.6
Zero economies	260.5	369.7	352.7	385.1
	Panel C: Difference (OA - Efficient)			
Baseline	15.8	32.6	16.8	44.9
Common economies	13.1	25.7	13.3	36.4
Zero economies	7.7	25.2	13.1	36.0

Note: Panels A and B present summary statistics for the observed or efficient allocation cost under alternative cost economies, holding fixed the simulated efficient allocations. The baseline case shows the same allocation cost as in table 4. The common economies rows set the volume and density cost parameters to their across-bidder and across-draw averages. The zero economies rows set these cost parameters to zero. All costs are in pesos per meal. As in the previous table, the Min-max column reports the results from the allocation that minimizes efficiency gains within the identified distribution-free bounds; the remaining columns are computed from the estimated type distribution.

potential complementarities and cost economies. While package bidding can theoretically take advantage of bidder-specific cost complementarities to achieve a low-cost allocation, the pass-through of these complementarities from costs to bids is somewhat limited in practice. This limitation is partly because a bidder's bid on a given package j competes with not only its competitors' bids, but also its own bid on other packages j'. As in a multi-product pricing problem in the presence of market power, a multi-package bidder internalizes the potential for "business stealing" by one of its bids from each of its other bids.

6 Conclusion

A large literature studies the identification of specific econometric models of behavior (see Matzkin, 2007 and the ensuing literature). Ensuring identification is essential in empirical analysis because it is a prerequisite for the validity of many statistical procedures (Newey and McFadden, 1994, for example). This paper develops a unified revealed-preference approach for identification in a class of models with a linear payoff structure. The approach is valid in a number of single-agent environments as well as in noncooperative games. It allows for both continuous and discrete actions, or a combination of the two. We characterize the identified set of models, show how to achieve point identification with excluded shifters, and use our results to suggest an estimation approach.

A number of important examples are a special case of our results. These include single-agent models where revealed preferences are used to identify preferences, for example, in multinomial choice models (McFadden, 1974, 1981); and reports made to strategy-proof mechanisms such as school choice mechanisms (Abdulkadiroglu et al., 2017) and second-price auctions. We also cover identification in games when the empirical strategy is based on implications of mutual best-responses in both private information settings (e.g., Guerre et al., 2000) as well as full-information cases (e.g., Berry et al., 1995).

These results also apply to cases beyond previously studied settings. As a case in point, our results apply immediately to recent extensions of scoring auctions studied in Asker and Cantillon (2008), including nonlinear multidimensional scoring auctions as in Hanazono et al. (2022) and scoring rules that combine discrete actions with continuous bids as in Aspelund and Russo (2023). These extensions are useful in a number of settings such as the decision of whether or not to provide one or more add-on services, which is a common feature of a number of auction settings.

Our results also suggest an estimation approach, which we illustrate by revisiting the com-

binatorial procurement auction for Chilean school lunches studied in Epstein et al. (2002) and KOW. Instead of directly parameterizing bidder markups, our approach targets the distribution of bidder costs and synergies during estimation. We estimate that in the 2003 combinatorial auction, the aggregate margin on awarded bids was at least 0.2 percent. However, a more efficient allocation was possible. We find that the allocation cost under the status-quo combinatorial auction exceeded that of the efficient allocation by at least 15.8 pesos per meal, or 3.8 percent of the total payment to winning bidders, and by 32.6 pesos per meal under the parametric cost distribution. In theory, package bidding in combinatorial auctions allows bidders to express across-unit cost complementarities. In practice, in Chile, we find that the pass-through of these complementarities from package costs to package bids is limited. The economically large efficiency gains arise from the ability of the efficient allocation to take full advantage of across-firm heterogeneity in economies of volume and density.

There are a number of issues that are left for future work. First, a strong restriction in our framework is that payoffs are linear in the outcome space, (x, t), which consists of an expected allocation and an expected transfer. Linear separability in outcomes and transfers rules out risk aversion. Relaxing this functional form is important. Second, estimation must confront the high dimensionality of the cost vector. We obtain distribution-free bounds on bidder types and on the cost of the efficient allocation, which we are able to compute after a suitable dimension reduction, and complement these with parametric point estimates. Some dimension reduction is difficult to avoid given the curse of dimensionality; relying on more flexible parametric cost distributions can be a feasible direction. Finally, we abstract away from issues such as endogeneity and unobserved heterogeneity, which are useful dimensions in which to extend our approach.

References

Abdulkadiroglu, Atila, Nikhil Agarwal, and Parag A. Pathak, "The Welfare Effects of Coordinated School Assignment: Evidence from the NYC High School Match," *American Economic Review*, 2017, *107*, 3635–89.

Agarwal, Nikhil and Paulo Somaini, "Demand Analysis Using Strategic Reports: An Application to a School Choice Mechanism," *Econometrica*, 2018, *86*, 391–444.

_ **and** _ , "Revealed Preference Analysis of School Choice Models," *Annual Review of Economics*, 8 2020, *12*, 471–501.

Allen, Roy and John Rehbeck, "Identification of Average Demand Models," *Econometrica*, 2017.

Artstein, Zvi, "Distributions of Random Sets and Random Selections," *Israel Journal of Mathematics*, 1983, *46* (4), 313–324.

Asker, John and Estelle Cantillon, "Properties of scoring auctions," *The RAND Journal of Economics*, 3 2008, *39*, 69–85.

Aspelund, Karl and Anna Russo, "Additionality and Asymmetric Information in Environmental Markets: Evidence from Conservation Auctions," *Working Paper*, 2023.

Athey, Susan and Philip A. Haile, "Identification of Standard Auction Models," *Econometrica*, 11 2002, *70*, 2107–2140.

_ **and** _ , "Nonparametric Approaches to Auctions," 1 2007.

Ausubel, Lawrence M. and Paul Milgrom, "The Lovely but Lonely Vickrey Auction," in Peter Cramton, Yoav Shoham, and Richard Steinberg, eds., *Combinatorial Auctions*, Cambridge, MA: MIT Press, 2006, pp. 17–40.

Bajari, Patrick and C.Â Lanier Benkard, "Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach," *Journal of Political Economy*, 12 2005, *113*, 1239–1276.

Beresteanu, Arie, Ilya Molchanov, and Francesca Molinari, "Sharp Identification Regions in Models with Convex Moment Predictions," *Econometrica*, 2011, *79* (6), 1785–1821.

Berry, Steven and Philip Haile, "Identification in Differentiated Products Markets," *Annual Review of Economics*, 10 2016, *8*, 27–52.

Berry, Steven T. and Philip A. Haile, "Identification in Differentiated Products Markets Using Market Level Data," *Econometrica*, 2014, *82*, 1749–1798.

_ **and** _ , "Identification of Nonparametric Simultaneous Equations Models With a Residual Index Structure," *Econometrica*, 2018, *86*, 289–315.

_ **, James Levinsohn, and Ariel Pakes**, "Automobile Prices in Market Equilibrium," *Econometrica*, 1995, *63*, 841 – 890.

Cantillon, Estelle and Martin Pesendorfer, "Combination Bidding in Multi-Unit Auctions," 2007.

Cassola, Nuno, Ali Hortaçsu, and Jakub Kastl, "The 2007 Subprime Market Crisis Through the Lens of European Central Bank Auctions for Short-Term Funds," *Econometrica*, 2013, *81* (4), 1309–1345.

Choquet, Gustave, Harry Corson, and Victor Klee, "Exposed points of convex sets," *Pacific Journal of Mathematics*, 1966, *17* (1), 33–43.

Ciliberto, Federico and Elie Tamer, "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 2009, *77*, 1791–1828.

Dieudonné, Jean, *Treatise on Analysis*, 1 ed., Vol. II, Academic Press, Inc., 1976.

Epstein, Rafael, Lysette Henríquez, Jaime Catalán, Gabriel Y. Weintraub, and Cristián Martínez, "A Combinational Auction Improves School Meals in Chile," *Interfaces*, 12 2002, *32*, 1–14.

Fu, Michael C., "Gradient Estimation," in Shane G. Henderson and Barry L. Nelson, eds., *Handbooks in Operations Research and Management Science*, Vol. 13 2006, pp. 575–616.

Galichon, Alfred and Marc Henry, "Set Identification in Models with Multiple Equilibria," *The Review of Economic Studies*, 2011, *78* (4), 1264–1298.

Gentry, Matthew L., Tatiana Komarova, and Pasquale Schiraldi, "Simultaneous First-Price Auctions with Preferences Over Combinations: Identification, Estimation and Application," *SSRN Electronic Journal*, 10 2014.

Gorman, W. M., "A Possible Procedure for Analysing Quality Differentials in the Egg Market," *The Review of Economic Studies*, 10 1980, *47*, 843–856.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong, "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 2000, *68*, 525–574.

Hanazono, Makoto, Yohsuke Hirose, Jun Nakabayashi, and Masanori Tsuruoka, "Theory and Estimation for Scoring Auctions," 11 2022.

Heitjan, Daniel F. and Donald B. Rubin, "Ignorability and Coarse Data," *The Annals of Statistics*, 1991, *19* (4), 2244–2253.

Kastl, Jakub, "Discrete Bids and Empirical Inference in Divisible Good Auctions," *The Review of Economic Studies*, 07 2011, *78* (3), 974–1014.

Kim, Sang Won, Marcelo Olivares, and Gabriel Y. Weintraub, "Measuring the Performance of Large-Scale Combinatorial Auctions: A Structural Estimation Approach," *Management Science*, 5 2014, *60*, 1180–1201.

Lancaster, Kelvin J., "A New Approach to Consumer Theory on JSTOR," *Journal of Political Economy*, 4 1966, *74*, 132–157.

Larsen, Bradley and Anthony Lee Zhang, "A Mechanism Design Approach to Identification and Estimation," 2018.

Lewbel, Arthur, "Semiparametric qualitative response model estimation with unknown heteroscedasticity or instrumental variables," *Journal of Econometrics*, 7 2000, *97*, 145–177.

_ **and Xun Tang**, "Identification and estimation of games with incomplete information using excluded regressors," *Journal of Econometrics*, 11 2015, *189*, 229–244.

Luo, Yao, Isabelle Perrigne, and Quang Vuong, "Structural Analysis of Nonlinear Pricing," *Journal of Political Economy*, 2018, *126*, 2523–2568.

Matzkin, Rosa L., "Nonparametric identification," 1 2007.

McCormick, Garth P., "Computability of Global Solutions to Factorable Nonconvex Programs: Part I — Convex Underestimating Problems," *Mathematical Programming*, December 1976, *10* (1), 147–175.

McFadden, Daniel, "Conditional Logit Analysis of Qualitative Choice Behavior," 1973, pp. 105 – 142.

_ , "Conditional logit analysis of qualitative choice behavior," *Frontiers in econometrics*, 1974, pp. 105–142.

_ , "Econometric models of probabilistic choice," *Structural analysis of discrete data with econometric applications*, 1981.

Molchanov, Ilya, *Theory of Random Sets*, 2 ed., London: Springer, 2017.

Newey, Whitney K. and Daniel L. McFadden, "Large Sample Estimation and Hypothesis Testing," in Robert F. Engle and Daniel L. McFadden, eds., *Handbook of Econometrics*, Vol. 4, Elsevier, 1994, pp. 2111–2245.

OECD, "Main Economic Indicators - complete database," 2016.

Pakes, A., "Alternative Models for Moment Inequalities," *Econometrica*, 2010, *78*, 1783–1822.

Rockafellar, R. Tyrrell, *Convex Analysis*, Princeton University Press, 1970.

Rosen, Sherwin, "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition on JSTOR," *Journal of Political Economy*, 1 1974, *82*, 34–55.

Samuelson, Paul A., "A Note on the Pure Theory of Consumer's Behaviour," *Economica*, 1938, *5*, 61–71.

Shabat, Boris Vladimirovich, *Introduction to Complex Analysis, Part 2*, American Mathematical Society, 1992.

Tebaldi, Pietro, Alexander Torgovitsky, and Hanbin Yang, "Nonparametric Estimates of Demand in the California Health Insurance Exchange," 5 2019.

Train, Kenneth E., "Discrete Choice Methods with Simulation," *Cambridge Books*, 2009.

Varian, Hal R., "The Nonparametric Approach to Demand Analysis," *Econometrica*, 1982, *50*, 945–973.

Xiao, Mo and Zhe Yuan, "License Complementarity and Package Bidding: The U.S Spectrum Auctions," 2020.

Appendix

A Lemmata

Lemma 1. *Let $\mathcal{X} \subset \mathbb{R}^J$ be nonempty and compact with $\bar{\mathcal{X}} := conv(\mathcal{X})$ of nonempty interior, and let $t = convt_{\mathcal{X}}$ be the convex hull of a lower semicontinuous function $t_{\mathcal{X}} : \mathcal{X} \to \mathbb{R}$. Assume that t is proper, closed, and convex with effective domain $\bar{\mathcal{X}}$. Then there exists $B \subset \mathcal{X}$ such that $\partial t(B) = \left\{v \in \mathbb{R}^J : v \in \partial t(x), x \in B\right\}$ has strictly positive Lebesgue measure in $\mathbb{R}^J$ and is contained in a translated simplicial cone.*

Proof. The conjugate function t^* is a proper closed convex function (Rockafellar, 1970, Theorem 12.2). For every $v \in \mathbb{R}^J$, $t^*(v) = \max_{x \in \bar{\mathcal{X}}} v \cdot x - t(x) = \max_{x \in \mathcal{X}} v \cdot x - t_{\mathcal{X}}(x)$; because t is proper and lower semicontinuous (Rockafellar, 1970, page 52) and $\bar{\mathcal{X}}$ is compact, this maximum is attained, so t^* is finite with effective domain $\mathbb{R}^J$ and $\partial t^*(v) \subset \bar{\mathcal{X}}$. The set D where t^* is differentiable is a dense subset of $\mathbb{R}^J$ of full Lebesgue measure, and the gradient mapping $\nabla t^* : v \to \nabla t^*(v)$ is continuous on D (Rockafellar, 1970, Theorem 25.5). At every $v \in D$ the maximizer $\nabla t^*(v)$ is unique, so the pair $(\nabla t^*(v), t(\nabla t^*(v)))$ is an exposed point of the convex hull of the epigraph of $t_{\mathcal{X}}$ and hence $\nabla t^*(v) \in \mathcal{X}$. Take any $v_0 \in D$ and let $x_0 = \nabla t^*(v_0) \in \mathcal{X}$. Let $S = \{x_1, x_2, ..., x_J\} \subset \mathcal{X}$ be such that the convex hull of $S \cup \{x_0\}$ has strictly positive Lebesgue measure in $\mathbb{R}^J$; such a set S exists because $\bar{\mathcal{X}}$ has nonempty interior, so $\mathcal{X}$ affinely spans $\mathbb{R}^J$. Let $\bar{x} \in \mathbb{R}^J$ be such that x_0 is in the interior of the convex hull of $S \cup \{\bar{x}\}$. Such an $\bar{x}$ exists, although it may not be an element of $\bar{\mathcal{X}}$. Let $\bar{B}(x_0)$ be a closed convex neighborhood of x_0 in the interior of the convex hull of $S \cup \{\bar{x}\}$. By the continuity of the gradient mapping ∇t^* on D, there is an open neighborhood $V(v_0) \subset \mathbb{R}^J$ such that $\nabla t^*(v) \in \bar{B}(x_0)$ for every $v \in V(v_0) \cap D$. Let $V_1(v_0)$ be a smaller open neighborhood of v_0 whose closure is contained in $V(v_0)$, and define $B = \bar{B}(x_0) \cap \mathcal{X}$. For every $v \in V_1(v_0) \cap D$, the maximizer $\nabla t^*(v)$ lies in $\bar{B}(x_0)$ and, by the above, in $\mathcal{X}$; hence $\nabla t^*(v) \in \bar{B}(x_0) \cap \mathcal{X} = B$, and therefore $v \in \partial t(\nabla t^*(v)) \subset \partial t(B)$ (Rockafellar, 1970, Theorem 23.5). Since D has full Lebesgue measure, $\partial t(B) \supset V_1(v_0) \cap D$ has strictly positive Lebesgue measure in $\mathbb{R}^J$.

Now, we show that $\partial t(B)$ is contained in a translated simplicial cone. Consider the choice set $\mathcal{Y} = S \cup \{\bar{x}\}$ with transfers $f_{\mathcal{Y}}(x) = t(x)$ for $x \in S$ and $f_{\mathcal{Y}}(\bar{x})$ sufficiently low so that all outcomes in B with transfers $t(\cdot)$ would be dominated by the outcomes in $\mathcal{Y}$. Because $\bar{B}(x_0)$ is compactly contained in the interior of the convex hull of $S \cup \{\bar{x}\}$, the barycentric weight placed on $\bar{x}$ is bounded away from zero uniformly over $x \in \bar{B}(x_0)$. Moreover, t is bounded below on the compact set B. Therefore, $f_{\mathcal{Y}}(\bar{x})$ can be chosen sufficiently low that,

for every $x \in B$, the lottery over $S \cup \{\bar{x}\}$ yielding x has a transfer strictly below $t(x)$. Let $f(\cdot)$ be the convex hull of the function $f_{\mathcal{Y}}(\cdot)$. The function $f(\cdot)$ is the largest convex function that lies below $f_{\mathcal{Y}}(\cdot)$ and coincides with it on the set $\mathcal{Y}$. The subdifferential $\partial f(\bar{x})$ is a translated simplicial cone because the set $\mathcal{Y}$ includes $J+1$ affinely independent points in $\mathbb{R}^J$. If $v \in \partial t(B)$, then an agent with value v weakly prefers some $x \in B$ to any $x' \in S$. However, by construction, dominance implies that some element of $S \cup \{\bar{x}\}$ is strictly preferred to x. Since the agent weakly prefers x to every element of S, this dominating element must be $\bar{x}$. Thus, the agent strictly prefers $\bar{x}$ to x and, by transitivity, to all $x' \in S$. Therefore, $\partial t(B) \subset \partial f(\bar{x})$ which is a translated simplicial cone. □

Lemma 2. *Let $C \subseteq \mathbb{R}^J$ be a Lebesgue-measurable set with strictly positive Lebesgue measure. Assume that $C \subseteq \bar{C} + \{x\}$, where $\bar{C}$ is a simplicial cone and $+$ denotes Minkowski summation. Then, there exists $\lambda \in int\, \mathcal{N}(\bar{C})$, such that*

$$\hat{\chi}_{C,\lambda}(\xi) = \int_C \exp(-2\pi v \cdot (i\xi + \lambda))\, \mathrm{d}v$$

is not zero on any open set $\Xi \subseteq \mathbb{R}^J$. Moreover, $\hat{\chi}_{C,\tilde{\lambda}}(\xi)$ is not zero on any open set $\Xi \subseteq \mathbb{R}^J$ for any $\tilde{\lambda} = \alpha\lambda$ for $\alpha \in (0,1)$.

Proof. We first introduce some notation. First, define the matrix $A_{\bar{C}} = \begin{bmatrix} a_1 & \cdots & a_J \end{bmatrix}$ so that $v \in \bar{C}$ if and only if $v = A_{\bar{C}} u$ for some $u \geq 0$. By definition, $A_{\bar{C}}$ is invertible because $\bar{C}$ is simplicial. It is without loss of generality to assume that $|\det(A_{\bar{C}})| = 1$. Second, let $\mathcal{N}\left(\bar{C}\right)$ be the dual cone to $\bar{C}$. Note that int $\mathcal{N}\left(\bar{C}\right)$ is nonempty because $\bar{C}$ is simplicial. Fix a $\lambda \in$ int $\mathcal{N}\left(\bar{C}\right)$ for the remainder of this proof.

Towards a contradiction, assume that $\hat{\chi}_{C,\lambda}(\xi)$ is zero on an open set $\Xi \subseteq \mathbb{R}^J$. We will show below that $\hat{\chi}_{C,\lambda}(\xi)$ when viewed as a function on $\Xi_\lambda = \{\xi \in \mathbb{C}^J : \xi = y + iz, y \in \mathbb{R}^J, \frac{\lambda}{2} - z \in \text{int}\, \mathcal{N}(\bar{C})\}$, which contains $\mathbb{R}^J$, is holomorphic. Under this hypothesis, Theorem 5 in Shabat (1992) implies that $\hat{\chi}_{C,\lambda}(\xi) = 0$ for all $\xi \in \mathbb{R}^J$. However, this contradicts the fact that $\hat{\chi}_{C,\lambda}(\xi)$ is, up to scale, the characteristic function of a random variable with density $1\{v \in C\} \exp(-2\pi \langle v, \lambda \rangle)$. This density, and therefore $\hat{\chi}_{C,\lambda}(\xi)$, is nonzero because C has strictly positive Lebesgue measure.

The second part follows immediately because if $\lambda \in$ int $\mathcal{N}\left(\bar{C}\right)$ then so is $\tilde{\lambda} = \alpha\lambda$ for any $\alpha \in (0,1)$.

Hence, it only remains to show that $\hat{\chi}_{C,\lambda}(\xi)$ is holomorphic in Ξ_λ. By the change of variables $w = v - x$, $\hat{\chi}_{C,\lambda}(\xi) = \exp(-2\pi x \cdot (i\xi + \lambda))\, \hat{\chi}_{C-x,\lambda}(\xi)$, where $C - x \subseteq \bar{C}$. The prefactor $\exp(-2\pi x \cdot (i\xi + \lambda))$ is entire in ξ and nowhere zero, so $\hat{\chi}_{C,\lambda}$ is holomorphic on Ξ_λ if and

only if $\hat{\chi}_{C-x,\lambda}$ is. Replacing C by $C-x$, we may therefore assume without loss of generality that $C \subseteq \bar{C}$. To do this, we first use the differentiation under the integral sign theorem for complex variables (Theorem 13.8.6(iii) in Dieudonné, 1976) to show that $\frac{\partial \hat{\chi}_{C,\lambda}(\xi)}{\partial \xi_k}$ exists on Ξ_λ and is equal to

$$\frac{\partial \hat{\chi}_{C,\lambda}(\xi)}{\partial \xi_k} = -2\pi i \int_C v_k \exp\left(-2\pi v \cdot (i\xi + \lambda)\right) \mathrm{d}v.$$

Fix $\xi_{-k} = (\xi_1, \ldots, \xi_{k-1}, \xi_{k+1}, \ldots, \xi_J)$ and define the function $f_C(v, \xi_k; \xi_{-k}) = 1\{v \in C\} \exp\left(-2\pi v \cdot (i\xi + \lambda)\right)$. To apply the result, we need to show that $f_C(v, \xi_k; \xi_{-k})$ is (i) analytic in ξ_k for almost all v, (ii) measurable in v for each ξ, and (iii) there exists an integrable function $g(v)$ such that for almost all ξ, $|f_C(v, \xi_k; \xi_{-k})| \leq g(v)$. Requirement (i) follows from the definition of $f_C(\cdot)$. Measurability is immediate given the definition of $f_C(v, \xi_k; \xi_{-k})$.

To show (iii) we will use the fact that $|f_{\bar{C}}(v, \xi_k; \xi_{-k})| \geq |f_C(v, \xi_k; \xi_{-k})|$ for all v and show that $|f_{\bar{C}}(v)| \leq g(v) = 1\left\{v \in \bar{C}\right\} \left|\exp\left(-2\pi v \cdot \frac{\lambda}{2}\right)\right|$ for all $\xi \in \Xi_\lambda$, where $g(v)$ is integrable. To this end, we first bound $|f_{\bar{C}}(v)|$ as follows:

$$\begin{aligned}
|f_{\bar{C}}(v)| &= 1\left\{v \in \bar{C}\right\} |\exp\left(-2\pi v \cdot (i\xi + \lambda)\right)| \\
&= 1\left\{v \in \bar{C}\right\} |\exp\left(-2\pi v \cdot (i(y + iz) + \lambda)\right)| \\
&= 1\left\{v \in \bar{C}\right\} \left|\exp\left(-2\pi v \cdot \left(\frac{\lambda}{2} + \left(\frac{\lambda}{2} - z\right)\right)\right)\right| \\
&\leq 1\left\{v \in \bar{C}\right\} \left|\exp\left(-2\pi v \cdot \frac{\lambda}{2}\right)\right| \equiv g(v).
\end{aligned}$$

The first equality uses the definition $\xi = y + iz$. The second equality follows the fact that

$$\begin{aligned}
|\exp\left(-2\pi v \cdot (i(y + iz) + \lambda)\right)| &= |\exp\left(-2\pi v \cdot iy\right)| \, |\exp\left(-2\pi v \cdot (\lambda - z)\right)| \\
&= |\exp\left(-2\pi v \cdot (\lambda - z)\right)|.
\end{aligned}$$

The third equality writes $\lambda - z = \frac{\lambda}{2} + \left(\frac{\lambda}{2} - z\right)$. The inequality uses the fact that $v \in \bar{C}$ and $\frac{\lambda}{2} - z \in \operatorname{int} \mathcal{N}(\bar{C})$, which imply $v \cdot \left(\frac{\lambda}{2} - z\right) > 0$. We now show that $g(v)$ is integrable.

$$\begin{aligned}\int g(v)\,\mathrm{d}v &= \int 1\left\{v \in \bar{C}\right\}\left|\exp\left(-2\pi v \cdot \frac{\lambda}{2}\right)\right|\mathrm{d}v \\ &= \int 1\left\{A_{\bar{C}}^{-1}v \geq 0\right\}\left|\exp\left(-2\pi v \cdot \frac{\lambda}{2}\right)\right|\mathrm{d}v \\ &= \int_{\mathbb{R}_+^J} \exp\left(-2\pi A_{\bar{C}}u \cdot \frac{\lambda}{2}\right)\mathrm{d}u.\end{aligned}$$

The first equality uses the definition of $A_{\bar{C}}$. The second substitutes $u = A_{\bar{C}}^{-1}v$, and uses the fact that $|\det(A_{\bar{C}})| = 1$ and $u \geq 0$.

It remains to show that $\int_{\mathbb{R}_+^J} \exp\left(-2\pi A_{\bar{C}}u \cdot \frac{\lambda}{2}\right)\mathrm{d}u$ is finite. Observe that $\frac{\lambda}{2} \in \operatorname{int}\mathcal{N}\left(\bar{C}\right)$. Re-write the integral as

$$\begin{aligned}&\int_{\mathbb{R}_+^J} \exp\left(-2\pi A_{\bar{C}}u \cdot \frac{\lambda}{2}\right)\mathrm{d}u \\ =&\int_{\mathbb{R}_+^J} \exp\left(-2\pi \|u\| A_{\bar{C}}\frac{u}{\|u\|} \cdot \frac{\lambda}{2}\right)\mathrm{d}u \\ =&\int_{\mathbb{S}_+^{J-1}}\int_0^\infty r^{J-1}\exp\left(-2\pi r A_{\bar{C}}u' \cdot \frac{\lambda}{2}\right)\mathrm{d}r\mathrm{d}\sigma(u') \\ =&\int_{\mathbb{S}_+^{J-1}} \frac{1}{(2\pi A_{\bar{C}}u' \cdot \frac{\lambda}{2})^J}\int_0^\infty s^{J-1}\exp(-s)\,\mathrm{d}s\mathrm{d}\sigma(u') \\ =&\int_{\mathbb{S}_+^{J-1}} \frac{1}{(2\pi A_{\bar{C}}u' \cdot \frac{\lambda}{2})^J}\Gamma(J)\,\mathrm{d}\sigma(u'),\end{aligned}$$

where $\sigma(u')$ is the proper surface measure of the sphere $\mathbb{S}^{J-1}$, $\mathbb{S}_+^{J-1} = \mathbb{S}^{J-1} \cap \mathbb{R}_+^J$, $u' = \frac{u}{\|u\|}$, and $\Gamma(\cdot)$ is the Gamma function. The first equality is trivial, the second from a change of variables $u' = \frac{u}{\|u\|}$ and $r = \|u\|$, the third from a change of variables $s = 2\pi r A_{\bar{C}}u' \cdot \frac{\lambda}{2}$ and the last from the definition of the Gamma function.

Because $\frac{\lambda}{2} \in \operatorname{int}\mathcal{N}\left(\bar{C}\right)$, we have that $(A_{\bar{C}}u' \cdot \frac{\lambda}{2})$ is strictly positive, with a strictly positive minimum $\kappa > 0$. Hence, we have that

$$\int_{\mathbb{S}_+^{J-1}} \frac{1}{(2\pi A_{\bar{C}}u' \cdot \frac{\lambda}{2})^J}\Gamma(J)\,\mathrm{d}\sigma(u') < (2\pi\kappa)^{-J}\Gamma(J)\int_{\mathbb{S}_+^{J-1}} 1\mathrm{d}\sigma(u').$$

Since $\sigma\left(\mathbb{S}^{J-1}\right)$ is finite, we have that the right hand side is finite.

Finally, Osgood's Lemma implies that $\hat{\chi}_{C,\lambda}(\xi)$ is holomorphic if it is continuous in ξ. For

any $h \in \mathbb{C}^J$,

$$|\hat{\chi}_{C,\lambda}(\xi + h) - \hat{\chi}_{C,\lambda}(\xi)| \leq \int 1\{v \in C\} |\exp(-2\pi v \cdot (i(\xi + h) + \lambda)) - \exp(-2\pi v \cdot (i\xi + \lambda))| \, \mathrm{d}v.$$

Since $\xi \in \Xi_\lambda$, there exists $\varepsilon > 0$ such that for $h \in \mathbb{C}^J$ with $\|h\| < \varepsilon$, $\xi + h \in \Xi_\lambda$. As both $\xi + h$ and ξ are in Ξ_λ, the integrand is dominated by $2g(v)$, which is integrable. And $1\{v \in C\} |\exp(-2\pi v \cdot (i(\xi + h) + \lambda)) - \exp(-2\pi v \cdot (i\xi + \lambda))| = 1\{v \in C\} |\exp(-2\pi v \cdot (i\xi + \lambda))| |\exp(-2$ 0, as $h \to 0$. Therefore, by the dominated convergence theorem, $|\hat{\chi}_{C,\lambda}(\xi + h) - \hat{\chi}_{C,\lambda}(\xi)| \to 0$ as $h \to 0$. □

The technical arguments are a significant generalization of Theorem A.2 in Agarwal and Somaini (2018). This previous result, which is special to a school choice model, dealt with the case in which $\partial t(x)$ is a convex cone. When $\partial t(x)$ is a convex cone, the Fourier transform of $\chi_{\partial t(x),\lambda}(u)$ can be computed in closed form, circumventing the need for more general arguments.

Proof of Theorem 1. Lemma 1 implies that there exists a set $A \subseteq \mathcal{X}$ whose image $C := \cup_{x \in A} \partial t(x)$ has strictly positive Lebesgue measure and is contained in a translated simplicial cone. Let $\mathcal{F}_k \subseteq \mathbb{L}^1\left(\mathbb{R}^J\right)$ be the space of functions that satisfy the integrability condition in Assumption 2 for a given value of k. Consider any two candidate densities f_1 and f_2 satisfying Assumption 2, and choose $k > 0$ small enough that both belong to $\mathcal{F}_k$. With a slight abuse of notation, define the operator $\Lambda_A : \mathcal{F}_k \to \mathbb{L}^\infty\left(\mathbb{R}^J\right)$ as

$$\Lambda_A[f](z) = \int 1\{u + z \in \cup_{x \in A} \partial t(x)\} f(u) \, \mathrm{d}u.$$

Each candidate satisfies $\Lambda_A[f_i](z) = P_{X|z}(A)$ for F_Z-almost every z. For every $f \in \mathcal{F}_k$, the map $z \mapsto \Lambda_A[f](z)$ is continuous because translations are continuous in $\mathbb{L}^1$; since Z has full support, equality F_Z-almost surely implies equality for every z. Setting $f = f_1 - f_2$, we have $\Lambda_A[f] = 0$ for every z, so it is enough to show that $\Lambda_A[f] = 0$ implies $f = 0$ almost everywhere.

Towards a contradiction, suppose $\Lambda_A[f] = 0$ and f is nonzero on a set with positive Lebesgue measure. Because C is contained in a translated simplicial cone, write $C \subseteq \bar{C} + \{x_0\}$ with $\bar{C}$ a simplicial cone. Let $\mathcal{N}\left(\bar{C}\right)$ be the dual cone to $\bar{C}$. By the definition of $\mathcal{F}_k$, there exists $\lambda \in \operatorname{int} \mathcal{N}\left(\bar{C}\right)$ sufficiently small so that $\exp(2\pi u \cdot \lambda) f(u)$ is integrable for all $f \in \mathcal{F}_k$. Fix

one such value of λ. Rewrite

$$\begin{aligned}\Lambda_A[f](z) &= \int 1\{u+z\in C\} f(u)\,\mathrm{d}u \\ &= \int 1\{u\in C\} f(u-z)\,\mathrm{d}u \\ &= \exp(2\pi z\cdot\lambda)\int 1\{u\in C\}\exp(-2\pi u\cdot\lambda)\exp(2\pi(u-z)\cdot\lambda) f(u-z)\,\mathrm{d}u.\end{aligned}$$

Since $\exp(2\pi z\cdot\lambda) > 0$ a.e., $\Lambda_A[f] = 0$ a.e. $\iff$ $\hat{\chi}_{C,\lambda}(\xi)\cdot\bar{\hat{f}}_\lambda(\xi) = 0$, where $\bar{\hat{f}}_\lambda$ is the conjugate of the Fourier transform of $f_\lambda(u) = \exp(2\pi u\cdot\lambda) f(u)$ and $\hat{\chi}_{C,\lambda}$ is the Fourier transform of $\chi_{C,\lambda} = 1\{u\in C\}\exp(-2\pi u\cdot\lambda)$. Since $\bar{\hat{f}}_\lambda(\xi)$ is continuous, the set of values ξ where $\bar{\hat{f}}_\lambda(\xi)\neq 0$ is open. Further, since f_λ is not zero almost everywhere, uniqueness of the Fourier transform implies that $\bar{\hat{f}}_\lambda(\xi)$ is not identically zero, so its support is nonempty. Therefore, there is a nonempty open set Ξ, such that $\bar{\hat{f}}_\lambda(\xi)\neq 0$ for all $\xi\in\Xi$. Because $\hat{\chi}_{C,\lambda}(\xi)\cdot\bar{\hat{f}}_\lambda(\xi) = 0$, it must be that for all $\xi\in\Xi$, $\hat{\chi}_{C,\lambda}(\xi) = 0$. However, this conclusion contradicts Lemma 2, which shows that $\hat{\chi}_{C,\lambda}$ is not zero on any open set $\Xi\subseteq\mathbb{R}^J$.

Let d be the dimension of $L^\perp$ and let T be an orthonormal transformation of outcomes and values that maps L onto the first $J-d$ coordinates and $L^\perp$ onto the last d. Such a T exists because L and $L^\perp$ are orthogonal complements. If $t(x)$ is differentiable at every $x\in\bar{\mathcal{X}}$ in every direction orthogonal to L, then the resulting $\tilde{t}(Tx) = t(x)$ is differentiable in the last d arguments, and z varies only in the first $J-d$. Thus, it is without loss of generality to assume that $t(x)$ is differentiable in the last d arguments. Therefore, the subdifferential $\partial t(x)$ is the Cartesian product of a set $V_1\subseteq\mathbb{R}^{J-d}$ and a point $v_2\in\mathbb{R}^d$. Let U_2 denote the last d elements of U. The observed choice x identifies the corresponding point v_2 and hence the last d elements of v. Since $u = v - z$ and z is observed, $f_{U_2}(\cdot)$ is identified. Consider the distribution of U_1, defined as the first $J-d$ elements of U, conditional on z and $U_2 = u_2$ in the support of U_2. Let $z = [z_1, z_2]$, where z_1 has full support in $\mathbb{R}^{J-d}$ and $z_2\in\mathbb{R}^d$ that, for simplicity is assumed to take a constant value of zero.

Let $\bar{\mathcal{X}}_1\subseteq\mathbb{R}^{J-d}$ be the projection of $\bar{\mathcal{X}}\subseteq\mathbb{R}^J$ onto its first $J-d$ dimensions. Consider the problem of an agent with type $v = [v_1, u_2]$ of choosing sequentially the optimal $x_1\in\bar{\mathcal{X}}_1$ first followed by a choice of x_2 so that $[x_1, x_2]\in\bar{\mathcal{X}}$. Given a choice of x_1, the agent's problem is to choose x_2 to maximize $u_2x_2 - t_2(x_2|x_1)$, where $t_2(x_2|x_1) = t([x_1, x_2])$. The conjugate function $t_2^*(u_2|x_1)$ denotes the maximum utility that an agent of type v can achieve in the second step if it chooses x_1 in the first one. Now, consider the first problem. The agent solves $\max_{x_1\in\bar{\mathcal{X}}_1} v_1\cdot x_1 - t_1(x_1|u_2)$, where $t_1(x_1|u_2) = -t_2^*(u_2|x_1)$. As a function of its first argument, $t_1(\cdot|u_2)$ is a lower semicontinuous convex function defined on a nonempty convex

and compact set $\bar{\mathcal{X}}_1$ by Assumption 1. $f_{U_1|u_2}(\cdot)$ satisfies Assumption 2 for almost all u_2. Because $\bar{\mathcal{X}}$ has nonempty interior, its projection $\bar{\mathcal{X}}_1$ has nonempty interior in $\mathbb{R}^{J-d}$. Thus, taking $\bar{\mathcal{X}}_1$ as the reduced choice set, Assumption 3 is satisfied. By part (i) of this theorem, $f_{U_1|u_2}(\cdot)$ is identified for almost all u_2. Therefore, the joint distribution of U is identified. □

Lemma 3. $\mathcal{R}_H = \mathcal{O}_H$.

Proof. Suppose $F_{V,Z} \notin \mathcal{R}_H$. Then there exist z and a Borel set $\mathcal{X}' \subseteq \mathcal{X}$ such that

$$P_{X|z}(\mathcal{X}') > \int 1\{v \in \cup_{x\in\mathcal{X}'}\partial t(x)\}\,\mathrm{d}F_{V|Z=z},$$

or equivalently,

$$P_{X|z}\left((\mathcal{X}')^c\right) < \int 1\{v \notin \cup_{x\in\mathcal{X}'}\partial t(x)\}\,\mathrm{d}F_{V|Z=z}.$$

Let $\mathcal{V} := \{\cup_{x\in\mathcal{X}'}\partial t(x)\}^c$. If $x \in \cup_{v\in\mathcal{V}}\partial t^*(v)$, then $v \in \mathcal{V}$ and $x \in \partial t^*(v)$ for some v, so $v \in \partial t(x)$ by the inverse-subgradient relation. Were $x \in \mathcal{X}'$, we would have $v \in \cup_{x\in\mathcal{X}'}\partial t(x)$, contradicting $v \in \mathcal{V}$. Hence $\cup_{v\in\mathcal{V}}\partial t^*(v) \subseteq (\mathcal{X}')^c$, and therefore

$$P_{X|z}(\cup_{v\in\mathcal{V}}\partial t^*(v)) \le P_{X|z}\left((\mathcal{X}')^c\right) < \int_{\mathcal{V}} \mathrm{d}F_{V|Z=z},$$

so $F_{V,Z} \notin \mathcal{O}_H$. Thus $\mathcal{O}_H \subseteq \mathcal{R}_H$.

Conversely, suppose $F_{V,Z} \notin \mathcal{O}_H$. Then there exist z and a Borel set $\mathcal{V} \subseteq \mathbb{R}^J$ such that

$$\int_{\mathcal{V}} \mathrm{d}F_{V|Z=z} > P_{X|z}(\cup_{v\in\mathcal{V}}\partial t^*(v)),$$

or equivalently,

$$\int_{\mathcal{V}^c} \mathrm{d}F_{V|Z=z} < P_{X|z}(\{\cup_{v\in\mathcal{V}}\partial t^*(v)\}^c).$$

Let $\mathcal{X}' := \{\cup_{v\in\mathcal{V}}\partial t^*(v)\}^c$. If $v \in \cup_{x\in\mathcal{X}'}\partial t(x)$, then $x \in \mathcal{X}'$ and $v \in \partial t(x)$ for some x, so $x \in \partial t^*(v)$. Were $v \in \mathcal{V}$, we would have $x \in \cup_{v\in\mathcal{V}}\partial t^*(v)$, contradicting $x \in \mathcal{X}'$. Hence $\cup_{x\in\mathcal{X}'}\partial t(x) \subseteq \mathcal{V}^c$, and therefore

$$\int 1\{v \in \cup_{x\in\mathcal{X}'}\partial t(x)\}\,\mathrm{d}F_{V|Z=z} \le \int_{\mathcal{V}^c} \mathrm{d}F_{V|Z=z} < P_{X|z}(\mathcal{X}'),$$

so $F_{V,Z} \notin \mathcal{R}_H$. Thus $\mathcal{R}_H \subseteq \mathcal{O}_H$. □

Proof of Proposition 2. If $P_{X|z}(\mathcal{X}^r) < 1$ for some z, then $\mathcal{R}_H = \mathcal{O}_H = \emptyset$ and $\mathcal{F}_H = \emptyset$. Now suppose $P_{X|z}(\mathcal{X}^r) = 1$ for all z.

Necessity ($\mathcal{F}_H \subseteq \mathcal{R}_H$). Let $K_{X^*|V,Z} \in \mathbb{K}$ reproduce the choice probabilities for $F_{V,Z} \in$

$\mathcal{F}_H$. For each v, $K_{X^*|V=v}$ is supported on $\partial t^*(v) \cap \mathcal{X}^r$, so $\int 1\{X^* \in \mathcal{X}'\} \mathrm{d}K_{X^*|V=v} \leq 1\{\partial t^*(v) \cap \mathcal{X}' \neq \emptyset\}$. By the inverse-subgradient relation $x \in \partial t^*(v) \Leftrightarrow v \in \partial t(x)$, the event $\partial t^*(v) \cap \mathcal{X}' \neq \emptyset$ coincides with $v \in \cup_{x \in \mathcal{X}'} \partial t(x)$; integrating over v gives $P_{X|z}(\mathcal{X}') \leq \int 1\{v \in \cup_{x \in \mathcal{X}'} \partial t(x)\} \mathrm{d}F_{V|Z=z}$, and $P_{X|z}(\mathcal{X}^r) = 1$ since $K_{X^*|V=v}$ is supported on $\mathcal{X}^r$. So $F_{V,Z} \in \mathcal{R}_H$.

Sufficiency ($\mathcal{R}_H \subseteq \mathcal{F}_H$). By Proposition 1(i), $x \in \mathcal{X}$ is optimal for v if and only if $t_{\mathcal{X}}(x) = t(x)$ and $v \in \partial t(x)$, so optimal outcomes are $\partial t^*(v) \cap \mathcal{X}^r$. Under Assumption 1(i)--(ii), this set is nonempty and compact for each v. For closed $\mathcal{X}' \subseteq \mathcal{X}$ and $F_{V,Z} \in \mathcal{R}_H$,

$$\begin{aligned} P_{X|z}(\mathcal{X}') = P_{X|z}(\mathcal{X}' \cap \mathcal{X}^r) &\leq \int 1\{v \in \cup_{x \in \mathcal{X}' \cap \mathcal{X}^r} \partial t(x)\} \mathrm{d}F_{V|Z=z} \\ &= \int 1\{v : \partial t^*(v) \cap \mathcal{X}^r \cap \mathcal{X}' \neq \emptyset\} \mathrm{d}F_{V|Z=z}. \end{aligned}$$

The first equality uses $P_{X|z}(\mathcal{X}^r) = 1$. The inequality uses membership in $\mathcal{R}_H$. The last equality uses the inverse-subgradient relation. By Artstein (1983), Theorem 3.1,[25] the inequalities above and compactness of $\partial t^*(v) \cap \mathcal{X}^r$ imply the existence of a measurable family $\mu(v) \in \Delta(\mathcal{X})$ with $\operatorname{supp} \mu(v) \subseteq \partial t^*(v) \cap \mathcal{X}^r$ such that

$$P_{X|z}(\mathcal{X}') = \int \mu(v)(\mathcal{X}') \mathrm{d}F_{V|Z=z}(v)$$

for all closed $\mathcal{X}' \subseteq \mathcal{X}$. Define $K_{X^*|V=v,Z=z} := \mu(v)$. Then $K_{X^*|V=v,Z=z}$ is supported on $\partial t^*(v) \cap \mathcal{X}^r$ for $F_{V|Z=z}$-almost every v, so $K_{X^*|V,Z} \in \mathbb{K}$ reproduces the choice probabilities and $F_{V,Z} \in \mathcal{F}_H$. Combined with necessity, $\mathcal{F}_H = \mathcal{R}_H$. By Lemma 3, $\mathcal{O}_H = \mathcal{R}_H$, so $\mathcal{F}_H = \mathcal{O}_H$. Finally, the restrictions define a convex set in the ambient space of distributions; consequently $\mathcal{F}_H$ is convex whenever $\mathbb{F}$ is convex. □

Proof of Proposition 3. Define

$$\widetilde{W}(q) = \mathbb{E}\left[\max_{x \in \mathcal{X}} \{x \cdot (U+q) - t(x)\}\right].$$

This expectation is finite. Let $R = \sup_{x \in \mathcal{X}} \|x\|$, fix $x_0 \in \mathcal{X}$, and let $\bar{t} = \inf_{x \in \mathcal{X}} t(x) > -\infty$. Then

$$-R\|u+q\| - t(x_0) \leq \max_{x \in \mathcal{X}} \{x \cdot (u+q) - t(x)\} \leq R\|u+q\| - \bar{t}.$$

Compactness of $\mathcal{X}$ and $\mathbb{E}\|U\| < \infty$ imply integrability. By the measurable maximum theo-

[25] In Artstein's notation, take $G = \mathbb{R}^J$, $B = \mathcal{X}$, sections $F(v) = \partial t^*(v) \cap \mathcal{X}^r$, and identify σ and ρ with $F_{V|Z=z}$ and the conditional law of X given z.

rem,

$$\widetilde{W}(q) = \max_{X:\mathbb{R}^J \to \mathcal{X}} \int [X(u) \cdot (u+q) - t(X(u))] \, \mathrm{d}F_U(u).$$

Let $m(z) = \mathbb{E}[X|z]$ and define $W(z) = \widetilde{W}(g(z))$. Optimality of the observed choice implies that, for every q,

$$\widetilde{W}(q) \geq \widetilde{W}(g(z)) + m(z) \cdot (q - g(z)),$$

and hence $m(z) \in \partial \widetilde{W}(g(z))$. Therefore, for any z',

$$m(z) \cdot (g(z') - g(z)) \leq W(z') - W(z) \leq m(z') \cdot (g(z') - g(z)).$$

By Assumptions 4(i)--(ii), g is continuously differentiable and m is continuous. The two bounds therefore imply

$$W(z') - W(z) = m(z) \cdot Dg(z)(z' - z) + o(\|z' - z\|).$$

Thus,

$$\nabla W(z) = Dg(z)^\top m(z),$$

and, because g is component-specific,

$$\frac{\partial W(z)}{\partial z_k} = \mathbb{E}[X_k|z] \, g_k'(z_k).$$

For $k \neq l$,

$$\frac{\partial^2 W(z)}{\partial z_l \partial z_k} = \frac{\partial \mathbb{E}[X_k|z]}{\partial z_l} g_k'(z_k).$$

By Assumptions 4(i)–(ii), these mixed partial derivatives exist and are continuous. Equality of mixed partial derivatives yields

$$\frac{\partial \mathbb{E}[X_k|z]}{\partial z_l} g_k'(z_k) = \frac{\partial \mathbb{E}[X_l|z]}{\partial z_k} g_l'(z_l),$$

and therefore

$$\frac{g_l'(z_l)}{g_k'(z_k)} = \frac{\partial \mathbb{E}[X_k|z] / \partial z_l}{\partial \mathbb{E}[X_l|z] / \partial z_k},$$

which is the desired equation. Therefore, for any value of g and z, and any pair k and l, we can identify

$$\frac{\partial g_l(z_l)}{\partial z_l} \Big/ \frac{\partial g_k(z_k)}{\partial z_k}. \tag{21}$$

The rest of the proof uses the arguments in Corollary S.5.1 in Allen and Rehbeck (2017).

First, we identify $\frac{\partial}{\partial z_j} g_j(z_0)$ up to scale. The first part of the proof established that $m(z) \in \partial \widetilde{W}(g(z))$. Since the subdifferential of a convex function is monotone,

$$(m(z) - m(z')) \cdot (g(z) - g(z')) \geq 0.$$

Now set $z = z_0$ and $z' = z_0 + \Delta e_j$, where $\Delta > 0$. Because g is component-specific, the preceding monotonicity inequality becomes

$$(\mathbb{E}[X_j | z_0] - \mathbb{E}[X_j | z_0 + \Delta e_j])(g_j(z_{0,j}) - g_j(z_{0,j} + \Delta)) \geq 0.$$

By Assumption 4(ii), the first factor is nonzero for sufficiently small Δ, and by Assumption 4(i), the second factor is nonzero. The inequality is therefore strict, and the sign of $g_j(z_{0,j}) - g_j(z_{0,j} + \Delta)$ is identified. Taking the limit of this difference divided by Δ as $\Delta \to 0$ implies that the sign of $\frac{\partial g_j(z_{0,j})}{\partial z_j}$ is identified for all j. The scale normalization implies that the partial derivatives of g are known at z_0.

Identification of $\frac{\partial g_j(z_{0,j})}{\partial z_j}$ and the ratios in equation (21) implies that for all l, $\frac{\partial g_l(z)}{\partial z_l}$ is identified for all z in its support. To show that this is sufficient to identify $g(z)$ up to scale and location, assume that there are two functions $g(z)$ and $\tilde{g}(z)$ that have the same partial derivatives and let $\delta(z) = g(z) - \tilde{g}(z)$. Since $\nabla \delta(z) = \nabla g(z) - \nabla \tilde{g}(z) = 0$ and the support of z is rectangular, we have that δ is the constant function. The location normalization implies that g is identified. □

B Empirical Appendix

B.1 Combinatorial Auction Mechanism

In this section, we present the full winner determination problem, as presented in KOW's appendix G. The decision variables $\tilde{x}_{ij} \in \{0, 1\}$ indicate whether bidder i is allocated package j. The combinatorial auction minimizes total payments (the sum of winning bids) by solving the integer program

$$
\begin{aligned}
\min_{\tilde{x}\in\{0,1\}^{IJ}} \quad & \sum_i \sum_j \tilde{x}_{ij} a_{ij} \\
\text{s.t.} \quad & \sum_i \sum_{j:k\in j} \tilde{x}_{ij} \geq 1 \quad \text{for all } k \\
& \sum_j \tilde{x}_{ij} \leq 1 \quad \text{for all } i \\
& \sum_j \tilde{x}_{ij} q_{ij} \leq \bar{q}_i \quad \text{for all } i \\
& \sum_j \tilde{x}_{ij} |j| \leq \bar{n}_i \quad \text{for all } i \\
& \underline{I}_r \leq \sum_i \sum_{j:j\cap r\neq\varnothing} \tilde{x}_{ij} \leq \bar{I}_r \quad \text{for all } r \\
& \sum_i \sum_j \tilde{x}_{ij} \geq \underline{I}
\end{aligned}
$$

In words, the constraints are:

1. Each unit k is assigned;
2. Each bidder i wins at most one package;
3. Each bidder i wins at most $\bar{q}_i$ total meals;
4. Each bidder i wins at most $\bar{n}_i$ total units;
5. Each region r is served by between $\underline{I}_r$ and $\bar{I}_r$ bidders; and
6. At least $\underline{I}$ total bidders are included in the allocation

In KOW, the authors write that the second constraint isn't explicitly imposed in the actual auctions, though it turns out that this constraint is not binding in most years. We follow KOW in including this second constraint. We never observe $\underline{I}$, the minimum number of bidders that must win in a given auction, needed to check the final constraint; we drop this final constraint in our application.

B.2 Bid Distribution

In tables B.1, B.2, and B.3, we present estimates $\hat{\theta}_a$ of the bid distribution parameters. In figure B.1, we compare the observed distribution of per-meal prices bid to our estimated distribution.

Table B.1: Estimates of Bid Adjustment Function Parameters β^{volume} and β^{density}

	Small bidders	Large bidders
Volume: Low	-5.545	-9.182
	(0.157)	(0.216)
Volume: Medium	-6.512	-12.165
	(0.187)	(0.219)
Volume: High	-11.320	-14.202
	(0.799)	(0.222)
Density: Medium	3.587	1.444
	(0.131)	(0.078)
Density: Low	5.065	1.914
	(0.161)	(0.094)

Note: Standard errors are in parentheses.

Table B.2: Estimates of Standard Deviations σ_ϵ of Idiosyncratic Bid Shocks

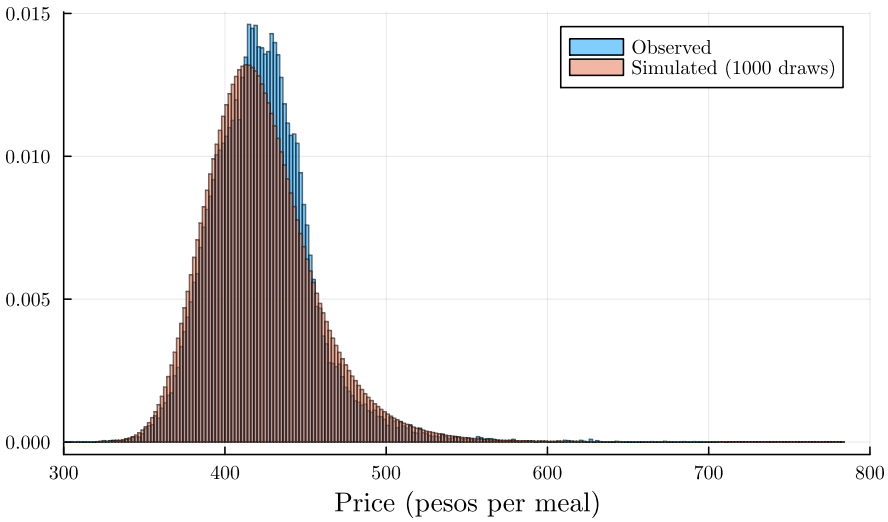
Package size (# units)	Standard deviation
1	14.324
2	5.050
3	4.041
4	3.122
5	2.429
6	1.923
7	1.247
8	1.454

Table B.3: Estimates of Means β^{unit} and Standard Deviations σ_ω of Unit Base Prices

	Mean	Standard deviation
Unit 401	485.968	14.053
Unit 402	427.033	11.843
Unit 403	501.982	14.175
Unit 404	592.235	33.294
Unit 405	566.486	32.240
Unit 501	487.112	28.038
Unit 502	407.773	17.705
Unit 503	417.131	15.238
Unit 504	403.168	20.036
Unit 505	420.122	11.507
Unit 506	432.474	13.872
Unit 507	437.179	12.005
Unit 508	390.913	14.702
Unit 510	402.816	8.991
Unit 511	421.090	29.663
Unit 901	496.375	25.829
Unit 902	497.620	24.353
Unit 903	572.034	35.821
Unit 904	429.400	17.323
Unit 905	472.411	23.963
Unit 906	459.874	19.570
Unit 907	485.291	23.961
Unit 908	513.643	25.055
Unit 909	566.889	38.081
Unit 1201	590.486	51.771
Unit 1331	401.659	13.172
Unit 1332	364.288	10.954
Unit 1333	342.243	11.160
Unit 1334	389.969	10.677
Unit 1335	378.911	21.624
Unit 1336	391.004	12.385
Unit 1339	430.929	20.292
Incumbency shifter	-18.998	

Note: The variance-covariance matrix Σ_ω is unrestricted, but for clarity only the (square roots of the) diagonal elements are reported here.

Figure B.1: Observed and Estimated Bid Distributions



Note: The blue histogram shows prices bid by all bidders on all packages on which they submitted bids in the data; an observation is a bidder-package pair. To construct the orange histogram, for each bidder-package pair, we simulate 1000 draws from the estimated bid distribution; each observation is a bidder-package-draw triple. Each histogram is normalized to represent a discrete probability distribution.

B.3 Type Bounds

To form the type bounds $\mathcal{G}_i$, we intersect the half-spaces implied by the revealed preference inequalities (15) for each deviating bid vector a_i'. We use one-quote deviations a_i' in which we change bidder i's quote on a single package $j \in \tilde{\mathcal{J}}_i$ and hold constant its quotes on other packages, $a_{i,-j}' = a_{i,-j}$. This appendix section describes how we construct the set $\tilde{\mathcal{J}}_i$ and choose the counterfactual quotes a_{ij}' for our downward and upward deviations.

The set $\tilde{\mathcal{J}}_i$ of deviating packages is constructed as follows. Among all packages that bidder i bids on in the data, we rank packages by their estimated probability of winning under the observed bid and set an initial threshold of 90% winning probability. We then iterate the following process: (i) consider packages with winning probabilities above the threshold and apply QR decomposition to the matrix of package-characteristic vectors in cost type space formed from this subset, leading to a basis of L' many linearly independent vectors with minimal collinearity; (ii) we stop the iteration if $L' = L$, where $L = \dim(\gamma)$, or the winning threshold is 0%; otherwise, we reduce the winning threshold by 0.1 percentage points and repeat step (i). Finally, we take the union of this basis with all single-unit packages that are feasible for any bidder.

Note that not all bidders are allowed to win packages involving all type dimensions. For example, some bidders are restricted in the size of packages they can win (where size is measured both in number of meals and number of units). If bidder i is not allowed to win any packages falling in the highest volume bin, then revealed preference can tell us nothing about the element of γ_i^{volume} corresponding to that volume bin.

Next, we compute, for each package $j \in \tilde{\mathcal{J}}_i$, the distribution of the **threshold quote** $\bar{p}_{ij}$, below which bidder i wins package j and above which bidder i loses that package, holding fixed its quotes on other packages. The uncertainty is over the bids of bidder i's opponents, which are random from i's perspective. Let $a_{-i}^{(s)}$ denote the vector of competitor bids in simulation draw s, and let $w^{(s)}$ denote the simulated procurement cost under the bids $\left(a_i, a_{-i}^{(s)}\right)$. For each bidder i, each package $j \in \tilde{\mathcal{J}}_i$ and simulation draw s:

1. If bidder i *loses* package j at $\left(a_i, a_{-i}^{(s)}\right)$, then we decrease its quote on that package from a_{ij} to p until it wins. To do this, let $w_{-i,-j}^{(s)}$ be the procurement cost when we exclude all of i's bids from eligibility and don't require that any units from package j be allocated. Bidder i wins package j at quote p when $p + w_{-i,-j}^{(s)} < w^{(s)}$, so the threshold quote is $\bar{p}_{ij}^{(s)} = w^{(s)} - w_{-i,-j}^{(s)}$.

2. If bidder i *wins* package j at $\left(a_i, a_{-i}^{(s)}\right)$, then we increase its quote on that package from a_{ij} to p until it loses. This time, let $\tilde{w}_{-ij}^{(s)}$ be the procurement cost when only package j from bidder i is removed from eligibility. Bidder i loses package j at quote p when $\tilde{w}_{-ij}^{(s)} < w^{(s)} - a_{ij} + p$, so the threshold quote is $\bar{p}_{ij}^{(s)} = \tilde{w}_{-ij}^{(s)} - w^{(s)} + a_{ij}$.

For each package $j \in \tilde{\mathcal{J}}_i$, the deviation quote a'_{ij} is selected from the distribution of these thresholds across simulation draws. For downward deviations, we use the 95th percentile of $\bar{p}_{ij}$ across draws in which bidder i loses package j under the observed bid, so that the deviation wins j in 5% of the draws in which bidder i loses j with the current bid. For upward deviations, we use the 50th percentile of $\bar{p}_{ij}$ across draws in which bidder i wins package j under the observed bid. For some bidders we instead use the 90th or 75th percentile for downward deviations, so that the deviation wins j in 10% or 25% of such draws, in order to avoid empty bounds.

B.4 Type Distribution

B.4.1 Gibbs Sampler

We sample from the posterior of $\theta_\gamma = (\mu_\gamma, \Sigma_\nu)$ using a Gibbs sampler with data augmentation. Collect the mean parameters into a vector μ_γ, consisting of volume and density means by bidder size, the unit-specific mean vector $(\mu_k^{\text{unit}})_k$, and the incumbency shifter μ^{incumb}. The type distribution can then be written

$$\gamma_i \mid z_i, \theta_\gamma \sim N\left(X_i \mu_\gamma, \Sigma_\nu\right),$$

where X_i is a known design matrix constructed from $z_i = (\text{size}_i, \text{incumb}_i)$.

We place conjugate, diffuse-but-proper priors on θ_γ:

$$\mu_\gamma \sim N\left(0, V_0\right), \qquad \Sigma_\nu \sim \mathcal{IW}\left(\nu_0, S_0\right),$$

with $V_0 = 10^{12} \cdot I_{n_\mu}$, where $n_\mu = \dim\left(\mu_\gamma\right)$ is the number of mean parameters, so that each mean parameter is assigned an independent prior with standard deviation 10^6 and the prior precision is negligible across any empirically relevant range. The inverse-Wishart hyperparameters are $\nu_0 = L + 2$ and $S_0 = 10{,}000 \cdot I_L$, the least informative proper inverse-Wishart prior given the dimension, with a diagonal scale that places independent, moderately diffuse mass on each variance.

Let N denote the number of bidder–auction observations. The sampler iterates three blocks.

Augmented types. For each i, draw γ_i from $N(X_i\mu_\gamma, \Sigma_\nu)$ truncated to $\mathcal{G}_i$. Because the cross-section of $\mathcal{G}_i$ in each dimension conditional on the others is an interval, this draw is taken via one-dimension-at-a-time Gibbs steps along the coordinates of γ_i.

Mean parameters. Conditional on $(\gamma_i)_i$ and Σ_ν, the augmented data form a multivariate Gaussian regression with closed-form normal posterior

$$\begin{aligned}\mu_\gamma \mid (\gamma_i)_i, \Sigma_\nu &\sim N(b_N, V_N),\\ V_N^{-1} &= V_0^{-1} + \sum_i X_i'\Sigma_\nu^{-1}X_i,\\ b_N &= V_N \sum_i X_i'\Sigma_\nu^{-1}\gamma_i.\end{aligned}$$

Covariance matrix. Conditional on $(\gamma_i)_i$ and μ_γ, the residuals are multivariate normal with mean zero, and

$$\Sigma_\nu \mid (\gamma_i)_i, \mu_\gamma \sim \mathcal{IW}\left(\nu_0 + N,\ S_0 + \sum_i (\gamma_i - X_i\mu_\gamma)(\gamma_i - X_i\mu_\gamma)'\right).$$

We run 4 independent chains for 200,000 iterations each. We discard the first 100,000 iterations of each chain as burn-in and thin the remainder by keeping every twentieth draw, yielding 5,000 draws per chain and $S_{\theta_\gamma} = 20{,}000$ total posterior draws of θ_γ. We report the posterior mean as our point estimate $\hat{\theta}_\gamma$ and the central 95 percent credible interval (2.5 and 97.5 percent marginal quantiles) for $\left(\mu^{\text{volume}}, \mu^{\text{density}}, (\mu_k^{\text{unit}})_k, \mu^{\text{incumb}}\right)$ and the diagonal elements of Σ_ν in tables 3 and B.4.

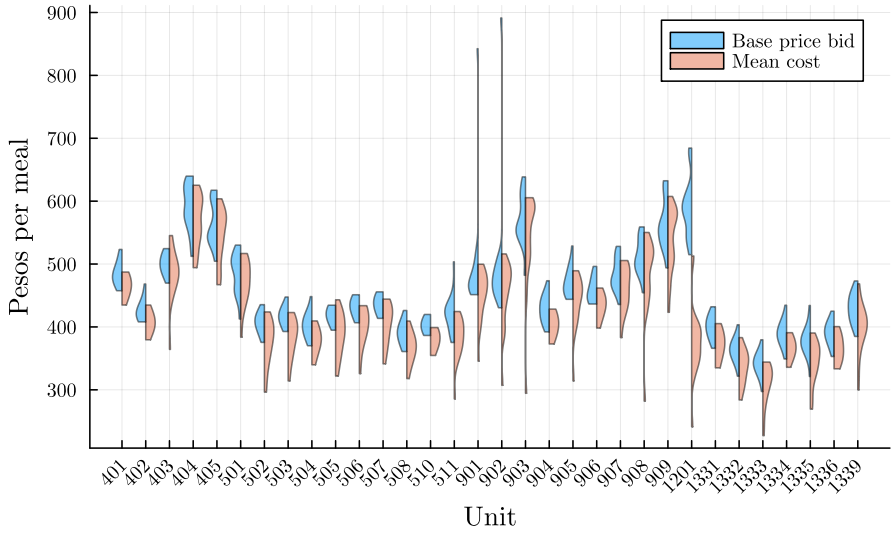
Table B.4: Gibbs Posterior Means of Unit Cost Parameters

	μ			σ		
		95% credible interval			95% credible interval	
	Mean	LB	UB	Mean	LB	UB
Unit 401	456.893	437.358	474.080	34.697	23.058	50.150
Unit 402	400.682	381.035	417.931	35.213	23.154	51.121
Unit 403	473.591	451.044	493.973	42.217	27.148	61.788
Unit 404	551.032	517.653	580.646	62.041	37.861	92.242

Table B.4: Gibbs Posterior Means of Unit Cost Parameters

	μ			σ		
		95% credible interval			95% credible interval	
	Mean	LB	UB	Mean	LB	UB
Unit 405	528.648	499.843	555.582	56.687	35.865	82.208
Unit 501	455.917	427.560	482.015	53.257	34.222	77.800
Unit 502	373.783	351.664	393.304	39.151	25.657	57.230
Unit 503	382.667	361.721	401.539	39.814	26.441	56.831
Unit 504	371.929	350.947	391.130	38.507	25.264	55.968
Unit 505	384.527	360.480	406.021	46.593	30.177	66.942
Unit 506	395.864	373.001	416.788	44.627	29.573	63.976
Unit 507	403.463	379.973	424.565	43.273	27.662	63.581
Unit 508	365.664	346.690	382.709	35.276	23.454	51.405
Unit 510	377.737	361.049	392.908	30.886	21.232	44.112
Unit 511	374.600	347.574	397.503	48.455	30.174	71.747
Unit 901	448.946	426.089	470.259	45.378	30.344	65.139
Unit 902	447.508	419.042	474.246	58.339	39.864	82.169
Unit 903	525.187	483.837	564.750	85.731	59.017	120.670
Unit 904	399.172	380.792	416.092	34.951	23.697	49.983
Unit 905	437.919	411.441	461.222	48.781	30.331	71.741
Unit 906	428.912	407.887	447.376	38.179	24.363	56.900
Unit 907	450.307	421.316	476.590	56.271	34.515	83.858
Unit 908	475.914	441.586	508.836	71.460	49.177	99.998
Unit 909	520.543	487.182	551.967	66.020	40.056	99.203
Unit 1201	482.684	434.879	520.964	59.029	26.630	109.502
Unit 1331	369.361	350.308	387.398	37.496	25.637	53.249
Unit 1332	333.057	314.054	350.012	35.593	24.240	50.907
Unit 1333	309.266	289.981	327.169	38.433	26.478	54.558
Unit 1334	359.947	342.083	376.747	34.845	23.974	49.305
Unit 1335	336.734	308.853	360.880	48.959	30.468	74.108
Unit 1336	360.973	341.460	379.309	38.472	25.916	54.649
Unit 1339	394.088	367.670	416.911	48.455	30.924	71.887
Incumbency shifter	-7.949	-21.200	5.253			

Figure B.2: Distributions of Bidders' Unit Base Prices and Costs



Note: These violin plots show the distributions of bidders' unit base prices $\tilde{\beta}_{ik}^{\text{unit}}$ (left, in blue) and unit costs $\gamma_{ik}^{\text{unit}}$ (right, in orange). In each unit k's violin plots, an observation is a bidder i. Bidder i's cost of supplying each unit k is computed from the mean of the estimated type distribution $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$ conditional on the type bounds $\mathcal{G}_i$.

Note: This table presents estimates of $\left(\left(\mu_k^{\text{unit}}\right)_k, \mu^{\text{incumb}}\right)$ and the corresponding diagonal elements of Σ_ν. The 95 percent credible interval (LB, UB) columns report the 2.5th and 97.5th percentiles of the marginal posterior distribution of each parameter.

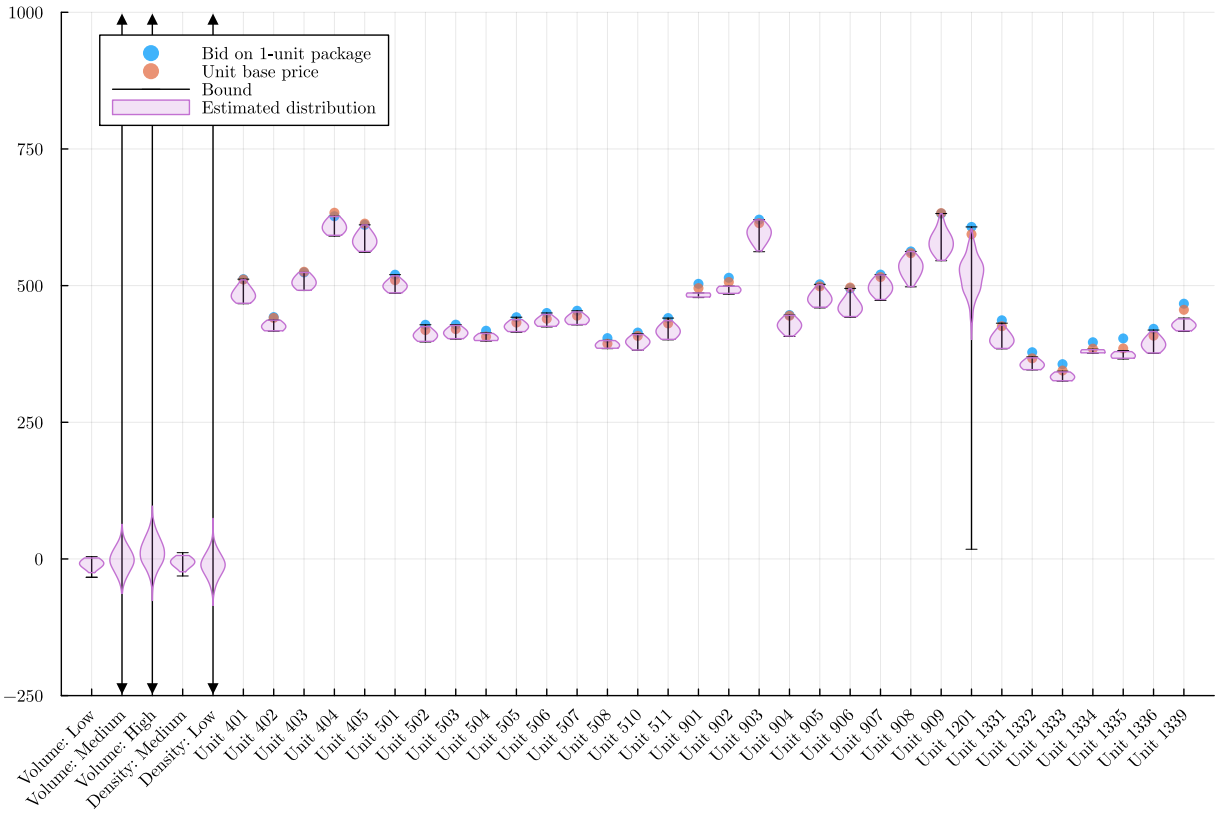
B.4.2 Type Distribution Estimates

Across all bidders and all units, the mean unit cost is 421 pesos per meal, compared to the mean unit base price of 450 pesos per meal discussed in section 5.3.2. The average margin from unit cost to unit base price is 7.5 percent of the unit base price. Units on which bidders bid higher base prices are also higher-cost to supply. Figure B.2 compares the two distributions separately for each unit.

B.5 Profit Margins

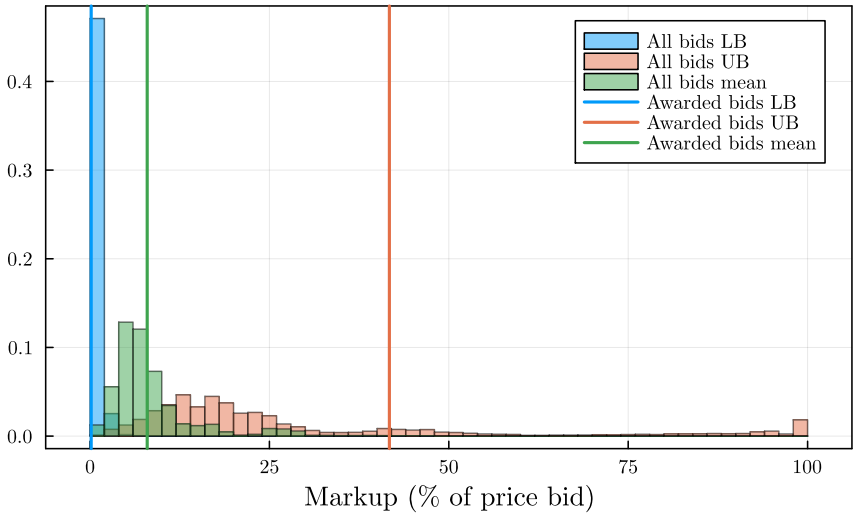
For each bid a_{ij} submitted by bidder i on package j, we compute the minimum and maximum profit margins $\frac{a_{ij}-c_{ij}}{a_{ij}} = \frac{a_{ij}-M_j\gamma_i}{a_{ij}}$ given the bounds on γ_i. Each package's margin bound is

Figure B.3: Example Bidder Type Estimates and Estimated Type Distribution



Note: This figure repeats figure 5 for the same example bidder i and adds violin plots showing the marginals of the estimated type distribution $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$.

Figure B.4: Distribution of Estimated Bid Profit Margins



Note: Each observation underlying the histogram is a bid that a bidder i submits on a package j. The blue and orange histograms show the distribution of lower and upper bounds, respectively, on bid margins. The green histogram shows the distribution of mean margins, as discussed in section 5.6. The blue, orange, and green vertical lines respectively indicate the lower bound, upper bound, and mean margins for the awarded allocation. The blue and orange series are distribution-free bounds; the green series comes from the estimated type distribution.

computed independently, so a different γ_i may attain each bound; we do not impose a single type across a bidder's packages. The distributions of these lower and upper bounds across bids are shown in blue and orange, respectively, in figure B.4. We bound the aggregate margin on awarded bids from below by 0.2 percent of the total payment to winning bidders and from above by 41.7 percent.

We compute expected profit margins using the estimated type distribution $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$ conditional on the type bounds $\mathcal{G}_i$. That is, for each bid a_{ij}, we compute $\mathbb{E}\left[\frac{a_{ij}-M_j\gamma_i}{a_{ij}} \mid \gamma_i \in \mathcal{G}_i, z_i; \hat{\theta}_\gamma\right]$. Figure B.4 plots the distribution of these expected margins across bidder-package pairs in green. The expected aggregate margin on awarded bids is 7.8 percent.

We present estimates of package costs and margins for each of the winning bids in table B.5.

Table B.5: Awarded Bid Cost and Profit Margin Estimates

Bidder	Units	Meals (millions)	Price bid (pesos per meal)	Cost (pesos per meal)			Markup (% of price bid)		
				LB	UB	Mean	LB	UB	Mean
10	[401, 402, 901, 902, 903, 906, 908, 909]	21.17	462.98	437.41	462.98	454.29	0.00	5.52	1.88
13	[905, 1334]	5.23	386.07	373.82	386.07	381.54	-0.00	3.17	1.17
16	[504, 505, 507]	6.42	391.58	32.85	391.58	348.43	0.00	91.61	11.02
17	[403, 404, 502, 503, 904, 907, 1201, 1331]	20.91	405.64	0.00	405.64	252.54	0.00	100.00	37.74
19	[1332, 1333, 1336]	9.48	340.05	293.53	340.05	321.56	0.00	13.68	5.44
26	[501]	1.64	459.07	425.85	442.61	435.96	3.59	7.24	5.03
28	[506, 508, 510, 511]	10.22	381.66	372.48	379.04	377.13	0.69	2.40	1.19
36	[1335, 1339]	3.60	381.77	0.00	381.77	285.62	-0.00	100.00	25.18
47	[405]	2.08	536.59	300.98	536.59	493.27	0.00	43.91	8.07

B.6 Efficiency

B.6.1 Bounds on Efficiency

Section 5.5 computes lower and upper bounds on the efficiency gains achievable by switching from the combinatorial auction to the efficient allocation. This appendix discusses technical details of solving the optimization problems (18) and (19). Throughout, we stack bidder-package pairs into vectors and write the objective compactly as $(M\gamma) \cdot (\tilde{x}^{OA} - \tilde{x})$.

To bound the efficiency gains from below, we must solve the saddle problem

$$\underline{\Delta C} = \min_{\gamma \in \mathcal{G}} \max_{\tilde{x} \in \tilde{\mathcal{X}}} (M\gamma) \cdot \left(\tilde{x}^{OA} - \tilde{x} \right). \tag{18}$$

When analyzing saddle problems, it is standard to invoke strong duality and transform the inner maximization problem into a minimization problem parameterized by γ. To dualize our problem, we need an explicit representation of the convex hull of $\tilde{\mathcal{X}}$. However, the constraints embedded in $\tilde{\mathcal{X}}$ break the equality between its convex hull and the set resulting from the relaxation of integrality constraints.

Instead, we work directly with the gain $\Delta C(\gamma)$ defined in (17). The definition of the maximum implies

$$\Delta C(\gamma) = \min_{t} \left\{ t : t \geq (M\gamma) \cdot \left(\tilde{x}^{OA} - \tilde{x} \right) \ \forall \ \tilde{x} \in \tilde{\mathcal{X}} \right\}.$$

This implies the following reformulation of the original problem (18):

$$\underline{\Delta C} = \min_{\gamma \in \mathcal{G},\, t \in \mathbb{R}} \left\{ t : t \geq (M\gamma) \cdot \left(\tilde{x}^{OA} - \tilde{x} \right) \ \forall \ \tilde{x} \in \tilde{\mathcal{X}} \right\}. \tag{22}$$

This is a linear program with finitely many linear inequality constraints, one for each feasible allocation $\tilde{x} \in \tilde{\mathcal{X}}$. Because $\tilde{\mathcal{X}}$ has a large number of elements, it is not practical to enumerate all constraints explicitly. We instead employ the following iterative constraint generation method to obtain an exact solution. Start by initializing $\text{LB}^{(0)} = -\infty$, $\text{UB}^{(0)} = +\infty$ and $\tilde{\mathcal{X}}^{(0)} = \varnothing$. In each iteration indexed by n:

1. Solve a relaxed version of the linear program (22) in which we replace $\tilde{\mathcal{X}}$ by $\tilde{\mathcal{X}}^{(n-1)}$. Denote the solution by $\left(t^{(n)}, \gamma^{(n)} \right)$ and the optimal value by $\text{LB}^{(n)}$. Since the relaxed problem has fewer constraints than the full problem, $\text{LB}^{(n)} \leq \underline{\Delta C}$.

2. Solve for the efficient allocation $\tilde{x}^{(n)}$ given the type profile $\gamma^{(n)}$. Compute the associated efficiency gains, $\Delta C^{(n)} = (M\gamma^{(n)}) \cdot \left(\tilde{x}^{OA} - \tilde{x}^{(n)} \right)$. Since $\Delta C^{(n)}$ is the optimal value of

the inner maximization in the saddle problem (18) at $\gamma^{(n)}$, $\Delta C^{(n)} \geq \underline{\Delta C}$. Update $\text{UB}^{(n)} = \min\left\{\text{UB}^{(n-1)}, \Delta C^{(n)}\right\}$.

3. If $\text{LB}^{(n)} = \text{UB}^{(n)}$, then both equal $\underline{\Delta C}$ and we can terminate. Otherwise, set $\tilde{\mathcal{X}}^{(n+1)} = \tilde{\mathcal{X}}^{(n)} \cup \left\{\tilde{x}^{(n)}\right\}$ and return to step 1.

Since $\tilde{\mathcal{X}}$ is finite, the algorithm is guaranteed to terminate.

To bound the efficiency gains from above, we would like to solve

$$\overline{\Delta C} = \max_{\gamma \in \mathcal{G},\, \tilde{x} \in \tilde{\mathcal{X}}} (M\gamma) \cdot \left(\tilde{x}^{OA} - \tilde{x}\right). \tag{19}$$

In this case, we have a direct optimization formulation, but the objective is bilinear and hence not globally concave. Algorithms that guarantee global optimality in the presence of integrality constraints and the number of dimensions in our environment are slow. Instead, we propose a fast alternative that bounds the maximum gains $\overline{\Delta C}$ in both directions.

1. **Lower bound on $\overline{\Delta C}$ (=: $\overline{\Delta C}^{BCA}$):** We use block coordinate ascent, exploiting the fact that the maximization problem (19) is an integer linear program when γ is fixed and a simple linear program when $\tilde{x}$ is fixed. Linearity means each can be solved efficiently.

   Initialize a feasible $\gamma^{(0)} \in \mathcal{G}$, and set $\Delta C^{(0)} = -\infty$. In each iteration indexed by n:

   (a) Solve for the efficient allocation $\tilde{x}^{(n)}$ given the type profile $\gamma^{(n-1)}$.

   (b) Solve for the type profile $\gamma^{(n)}$ that maximizes the gains at the efficient allocation $\tilde{x}^{(n)}$, and denote the optimal value by $\Delta C^{(n)}$. Since $\Delta C^{(n)}$ is the optimal value of the maximization problem (19) at a fixed $\tilde{x}$, $\Delta C^{(n)} \leq \overline{\Delta C}$.

   (c) If $\left|\Delta C^{(n)} - \Delta C^{(n-1)}\right|$ is below some tolerance, stop and return $\Delta C^{(n)}$. Otherwise, return to step a.

   Block coordinate ascent generates a nondecreasing sequence of objective values $\Delta C^{(n)}$, but the limit may not be a global optimum, since the objective function is not globally concave. For this reason, we run the algorithm from S diverse starting points and retain the best solution, which we denote $\overline{\Delta C}^{BCA}$.

   The initial points $\left\{\gamma^{(0,s)}\right\}_{s=1}^{S}$ are generated as extreme points of $\mathcal{G}$. First, obtain any feasible point $\gamma^{c} \in \mathcal{G}$. Then, for each start s,

(a) Draw a direction $d^{(s)}$: If $s = 1$, set $d^{(1)} = M^T \tilde{x}^{OA}$, the direction of highest cost at the observed allocation. Otherwise, sample a direction $d^{(s)}$ uniformly from the unit sphere in $\mathbb{R}^{\dim(\gamma)}$.

(b) Compute the extreme point of $\mathcal{G}$ in the direction $d^{(s)}$:

$$\gamma^{(0,s)} \in \underset{\gamma \in \mathcal{G} \cap B(\gamma^c, R)}{\operatorname{argmax}} \; d^{(s)} \cdot \gamma,$$

where $B(\gamma^c, R) = \{\gamma : \|\gamma - \gamma^c\|_\infty \leq R\}$ is a box trust region that ensures boundedness of $\gamma^{(0,s)}$ when $\mathcal{G}$ is unbounded in the direction $d^{(s)}$.

2. **Upper bound on $\overline{\Delta C}$ (=: $\overline{\Delta C}^{MC}$):** First, we compute finite lower and upper bounds on each component of $M\gamma$ over $\mathcal{G}$. Using these bounds, McCormick envelope linearization gives an exact mixed integer linear reformulation of the original maximization problem (19) (McCormick, 1976). The decision variables in the new problem are $\gamma \in \mathcal{G}$, $\tilde{x} \in \tilde{\mathcal{X}}$, and new auxiliary variables $w \in \mathbb{R}^{IJ}$. We then relax the integrality constraints on $\tilde{x}$. The optimal value of the relaxed problem, $\overline{\Delta C}^{MC}$, is an upper bound on the optimal value of the original problem, $\overline{\Delta C}$.

In practice, the lower and upper bounds for $\overline{\Delta C}$ lie within 5 percent of one another at every point of the frontier, where the gap is measured as $(\overline{\Delta C}^{MC} - \overline{\Delta C}^{BCA})/\overline{\Delta C}^{MC}$. Figure 6 plots $\overline{\Delta C}^{BCA}$, which is attained by a type profile and allocation feasible under the estimated bounds. The true maximum-gains frontier therefore lies weakly below the curve shown.

B.6.2 Joint Distribution of Allocation Costs

Figure 7 displays the joint distribution, implied by the estimated type distribution, of the per-meal cost of the observed allocation and the per-meal cost of the efficient allocation. Each draw of the bidders' cost types $\{\gamma_i\}$ from the conditional distributions described above maps to a pair $m = \left(c^{OA}, c^{EA}\right)$, where c^{OA} is the per-meal cost of the observed allocation and c^{EA} the per-meal cost of the cost-minimizing allocation at those types. These pairs lie in the band Q between the minimum- and maximum-gains frontiers, and the contours report a kernel estimate of the density of m on Q.

The economically interesting region lies near the frontiers, which the conditional draws sample only sparsely. To estimate the density there precisely, we use importance sampling: we draw from proposal distributions anchored at the vertices of the frontiers and reweight each draw by the ratio of the posterior density of γ to the proposal density, pooling the draws across

anchors against their mixture proposal. This concentrates draws in the relevant region while preserving consistency for the posterior density of m.

We then form a boundary-corrected kernel density estimate

$$\hat{f}_m(\xi) = \frac{1\{\xi \in Q\}}{\beta(\xi)} \sum_s \bar{w}_s K_H(m_s - \xi)$$
$$\beta(\xi) = \int_Q K_H(u - \xi)\, \mathrm{d}u,$$

where $\{m_s\}$ are the importance draws, $\bar{w}_s$ their normalized importance weights, and K_H a product kernel with bandwidth matrix H. The factor $\beta(\xi)$ renormalizes the kernel by its mass inside Q, which prevents the estimate from leaking density across the frontiers. Figure 7 plots $\log \hat{f}_m$, shaded so that darker regions correspond to more likely cost pairs.