Identification using Revealed Preferences in Linearly Separable Models*

Nikhil Agarwal, Pearl Z. Li and Paulo Somaini†

May 11, 2025

Abstract

Revealed preference arguments are commonly used when identifying models of both single-agent decisions and non-cooperative 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 choice sets. 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 if the choice set is sufficiently rich or if a variable that shifts preferences is available. Our identification results also suggests an estimation approach. Finally, we implement this approach to estimate values in a combinatorial procurement auction for school lunches in Chile. An efficient auction can reduce costs by taking greater advantage of firm-specific cost complementarities.

*We are grateful to Gabriel Weintraub for sharing the data used in this project. We thank Brent Hickman, Francesca Molinari, Parag Pathak, Anna Russo, and Whitney Newey for helpful comments. Yucheng Shang 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. Pearl Z. Li: Department of Economics, Stanford University. Somaini: Stanford Graduate School of Business and NBER, email: soma@stanford.edu.

1 Introduction

Revealed preference arguments 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 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 class of non-cooperative games. Single-agent examples include hedonic demand models (e.g., Rosen, 1974; Bajari and Benkard, 2005) and monopoly pricing. Examples of non-cooperative 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 common question 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

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

where $v \in \mathbb{R}^J$ is the agent's preference type, and $x_{\mathcal{A}}(\cdot)$ and $t_{\mathcal{A}}(\cdot)$ are functions that map actions to outcomes and transfers respectively. We assume that the analyst knows (or can identify) $x_{\mathcal{A}}(a)$ and $t_{\mathcal{A}}(a)$ and observes each agent's actions. An important requirement for conducting counterfactual analysis is that the cumulative distribution function (CDF) of the random variable v given observables z, denoted with $F_{V|Z}(v; z)$, is identified. In addition to this distribution, some cases also yield identification of the payoff type for each agent in the

[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).

market.

This paper studies the identification of models with this linearly separable structure. Our framework and results allow for both discrete and continous choice sets $\mathcal{X}$ or a mixture of the two, and can incorporate both single-agent decision settings as well as a large class of non-cooperative games with independent private information. As we formally demonstrate in section 3, 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 non-cooperative 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. Similarly, in the model of Lewbel and Tang (2015), $a \in \{0,1\}$ denotes a binary, x_A is the identity function, and $t_A(a) = a \times z$ where z is a special regressor.

The main results characterize the identified set of distributions $F_{V|Z}(v|z)$ and derives conditions under which it is point identified. Point identification can be achieved in two cases. The base case is if the choice set is $\mathcal{X}$ is sufficiently "rich" so that each choice is optimal only for a unique payoff type. 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 the 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 non-linear form by applying arguments similar to those in Allen and Rehbeck (2017). Our 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. While we do not directly aim to extend the analysis of identification in these models, we hope that the relatively sparse structure required for our results will be useful for new models. For instance, we allow for a combination of discrete and continuous choices, which might be useful in some contexts – e.g., Aspelund and Russo (2023), which analyzes a scoring auction with multi-dimensional bids that include discrete and continuous components. We hope that our general results and this shared structure will yield more immediate results for other models where identification results do not currently exist.

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, 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 solution to this dimensionality problem than the one taken in KOW.

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.

Our paper shares its focus on general revealed preference arguments with Pakes (2010), which also starts with revealed preference inequalities. The models and approaches are non-nested – Pakes (2010) allows for expectational errors in agents' beliefs but places stronger functional form restrictions on the estimand, which is the expectation of v and z. We study the identification of the conditional distribution given by $F_{V|Z}$. In this sense, our estimand is similar to that in Ciliberto and Tamer (2009), although we pursue a non-parametric approach and consider conditions for obtaining point identification.

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 non-cooperative games, and 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. Each action results in an **outcome** described by $x \in \mathcal{X} = \left\{x \in \mathbb{R}^J : x = x_{\mathcal{A}}(a) \;\; a \in \mathcal{A}\right\}$ 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 the specific choice set $(x_{\mathcal{A}}, t_{\mathcal{A}})$ faced by an agent.

Agent i's preference type is denoted $v_i \in \mathbb{R}^J$. Her (expected) utility from choosing $a \in \mathcal{A}$ is given by the linear form

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

We assume that each agent chooses $a \in \mathcal{A}$ to maximize this indirect utility.[3]

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 $x \in \mathcal{X}$:

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

where we use the convention that the infimum of an empty set is infinity. Observe that is it dominated to choose an action $a \in \mathcal{A}$ if $(x_{\mathcal{A}}(a), t_{\mathcal{A}}(a))$ is not in the graph of $t_{\mathcal{X}}(x)$.

We introduce two 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

[2] Our allows for the chosen action to result from a mixed strategy that randomizes between actions in the set $\mathcal{A}$.

[3] This formulation embeds scale and location normalizations because the expected utility is equal to $-t_{\mathcal{A}}(a)$ if $x_{\mathcal{A}}(a) = 0$.

the function (Rockafellar, 1970, page 36):

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

where conv epi $t(\cdot)$ is the convex hull of the epi-graph of $t(\cdot)$.[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 non-empty 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, $\nabla t(x)$.

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 co-ordinates $(x_A, t_A), \ldots, (x_F, t_F)$. 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.

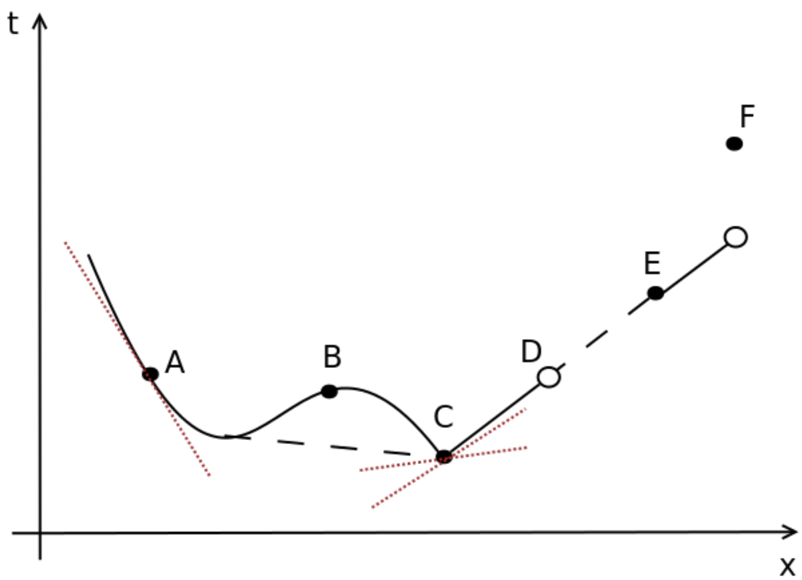
Since $t_{\mathcal{X}}(\cdot)$ and $t(\cdot)$ are differentiable at x_A, the only element of the subdifferential is the slope of the dotted line through A. In contrast, these functions are not differentiable at C. The subdifferentials of $t(\cdot)$ consists of the slopes of all lines that are everywhere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

2.2 Non-Cooperative Games

Although we began with the single-agent case, our framework also allows analysis of simultaneous move games that satisfy the following structure on payoffs and information. 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

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

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 outcomes. When an outcome is represented by a solid circle, it belongs to the graph of $t_{\mathcal{X}}(x)$ and results from an action $a \in \mathcal{A}$. An outcome represented by a hollow circle does not belong to the graph of $t_{\mathcal{X}}(x)$. Such points do not yield from an action $a \in \mathcal{A}$.

(a_i, a_{-i}). If the outcome is random conditional on the action profile, then interpret $\tilde{x}_i(a_i, a_{-i})$ and $\tilde{t}_i(a_i, a_{-i})$ be 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

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

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

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

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

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 Regularity

We will 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 $t_{\mathcal{X}}: \mathcal{X} \rightarrow \mathbb{R} \cup\{\infty\}$ is lower semi-continuous and takes finite values for some $x \in \mathcal{X}$.*

(ii) The set $\mathcal{X}$ is non-empty 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 maximimum is attained because Assumptions 1(i) and 1(ii) imply that $t(\cdot)$ is a lower semi-continuous proper convex function supported over a non-empty convex and compact set $\bar{\mathcal{X}}$, the convex hull of $\mathcal{X}$.

2.4 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. The analyst observes Z and the chosen outcome X.

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 $X = x_{\mathcal{A}}(A)$. In sum, we consider that the distribution of X conditional on Z is identified directly from the observed data.

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 explictly put together choices of agents with different values of the observables.

2.5 Identification

We follow the standard definitions of identification and falsifiability in the literature (e.g., Athey and Haile, 2002; Matzkin, 2007). That is, a model is identified if the joint distribution of the model's primitives is uniquely determined by the joint distribution of observables. In our case, 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, $F_{X,Z}$. We assume that the model is correctly specified. That is, the true value of $(F_{V,Z}, \phi)$ that generates the data belongs to $(\mathbb{F}, \mathbb{S})$.

Our analysis will focus on identifying various features of the model, such as the joint distribution $F_{V,Z}$:

Definition 1. A feature $\psi : (\mathbb{F}, \mathbb{S}) \to \Omega$ is **identified** given the model $(\mathbb{F}, \mathbb{S})$ if and only if for any two pairs $(F_{V,Z}, \phi)$ and $\left(\tilde{F}_{V,Z}, \tilde{\phi}\right)$ in $(\mathbb{F}, \mathbb{S})$, $\phi\left(F_{V,Z}\right) = \tilde{\phi}\left(\tilde{F}_{V,Z}\right)$ implies that $\psi\left(F_{V,Z}, \phi\right) = \psi\left(\tilde{F}_{V,Z}, \tilde{\phi}\right)$.

In addition to identification, some of our results will analyze whether the implications of a model are refutable:

Definition 2. A model $(\mathbb{F}, \mathbb{S})$ is **falsifiable** if and only if $\bigcup_{\phi \in \mathbb{S}, F_{V,Z} \in \mathbb{F}} \phi\left(F_{V,Z}\right) \subsetneq \mathbb{H}$.

Just as identification of a model in 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 $\mathcal{F}_I\left(H\right)$for observables H as the set of distributions $F_{V,Z}$ that are not refuted by H.

Definition 3. The identified set for a distribution of observables $H \in \mathbb{H}$ is

$$\mathcal{F}_I\left(H\right) = \left\{ F_{V,Z} \in \mathbb{F} : H \in \bigcup_{\phi \in \mathbb{S},} \phi\left(F_{V,Z}\right) \text{ and} F_Z = H_Z \right\},$$

where the restriction that $F_Z = H_Z$ indicates that the marginal distribution of Z is directly identified.

It will be useful to define $\mathbb{K}$ as the set of distributions of $X|V, Z$ that satisfy Assumption 1.(iii), i.e., where the distribution of $X|V, Z$ has support on the set of outcomes that are optimal of each v. Thus, if $\Pr\left(X \in \mathcal{X}'|z\right)$ denotes the probability of event $\mathcal{X}'$ conditional on z under $H \in \mathbb{H}$; the identified set is the set of distributions $F_{V,Z} \in \mathbb{F}$ such that there exist $K_{X^*|V,Z} \in \mathbb{K}$for which

$$\Pr\left(X \in \mathcal{X}'|z\right) = \int \left(\int 1\left\{X^* \in \mathcal{X}'\right\} \mathrm{d}K_{X^*|V=v} \right) \mathrm{d}F_{V|Z=z},$$

for all z and $\mathcal{X}' \subseteq \mathcal{X}$.

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

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

where $p\left(\cdot\right)$ is the pricing function, x_{kj} denotes the $j-$th characteristic of product k, and β_{ij} denotes the random co-efficients. This model fits our framework with $v_i = \left(\beta_{i1}, \ldots, \beta_{iJ}\right)$ and $t\left(\cdot\right) = p\left(\cdot\right)$.

This model is 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 $p\left(\cdot\right)$ and then using data on purchases to recover the marginal willingness to pay for x_{kj}, β_{ij}, between 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 applies to the identification of 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 $t\left(x\right)$ 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:

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

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 Unit, Independent Private Value (IPV) Auctions. Our model embeds standard IPV auctions that award the object to the highest bidder. In our notation, 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 $(v-a)\,H(a)$, where $H(a)$ 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 $H(a)$. The bid a parametrizes $x_{\mathcal{A}}(a) = H(a)$ and $t_{\mathcal{A}}(a) = aH(a)$.

We assume that all bids are observed. The identification of these auctions under weaker assumptions and 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 incoporate mechanisms more general than single-unit IPV auctions. 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. 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, $\tilde{x}\left(a_i, a_{-i}\right)$ be the quantity sold by firm i, and $\tilde{t}\left(a_i, a_{-i}\right)$ be the negative of the firm's revenue. Therefore, $x_{\mathcal{A}}\left(a_i; \mathcal{J}_i\right)$ and $t_{\mathcal{A}}\left(a_i; \mathcal{J}_i\right)$ are the expected quantities and the negative of expected revenues, respectively. Profit maximization implies that the firm maximizes

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

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 $x_i\left(a_i; \mathcal{J}_i\right) = a_i$. The function $t_i\left(a_i; \mathcal{J}_i\right)$ still denotes the expected negative revenue.

Example 6. School Choice. 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 $x_{\mathcal{A}}\left(a\right) \in \Delta^J \subseteq \mathbb{R}^J$, where $x_{\mathcal{A}}\left(a\right)$ is derived as in section 2.2. Let $d_i = \left(d_{i1}, \ldots, d_{iJ}\right)$ 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

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

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

Example 7. Trading Games. Larsen and Zhang (2018) consider a trading game in which an agent can take a sequence of actions $a = \left(a_1, \ldots, a_M\right)$. Following this sequence, the agent has a probability $x_{\mathcal{A}}\left(a\right) \in [0, 1]$ of engaging in a transaction and paying an expected (possibly negative) transfer $t_{\mathcal{A}}\left(a\right)$. 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

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

where the actions of other agents are integrated over as in section 2.2. This trading game is a one-dimensional special case of our model.

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

4 Revealed Preferences

4.1 Rationalizable Actions

Our first result derives a testable implication of the model and characterizes the consequences x that are optimal for some $v \in \mathbb{R}^J$.

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 $x \in \mathcal{X}$ 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, $t_{\mathcal{X}}(x) \geq t(x)$, 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 $\sum_j \alpha_j \left(v \cdot x_j - t_{\mathcal{X}}(x_j)\right) = v \cdot x' - t(x')$. Since $t_{\mathcal{X}}(x) = t(x)$, 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.

Part (ii). Under the maintained assumptions, $t_{\mathcal{X}}(x) = t(x)$ for all $x \in \mathcal{X}$ and $\partial t_{\mathcal{X}}(x) = \partial t(x)$. Moreover, $t(x)$ is continuous, implying that $\partial t(x)$ is non-empty. 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 $t(x)$ 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 an optimal choice. 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.

That said, part (ii) shows that if the cost function $t_{\mathcal{X}}(\cdot)$ is convex and smooth, then essentially any choice can be rationalized. Convexity implies that $t_{\mathcal{X}}(x) = t(x)$ for all $x \in \mathcal{X}$, ruling out points like B in Figure 1. Lower semi-continuity rules out points like D and F in the graph of $t_{\mathcal{X}}(\cdot)$, illustrated in Figure 1. The outcome $x_D \notin \mathcal{X}$ and therefore it does not belong to the domain of $t_{\mathcal{X}}(\cdot)$ although it belongs to the domain of $t(\cdot)$. 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 $\partial t(x)$ is not well-defined at this point. Together, convexity and lower semi-continuity imply that all solutions to the problem $\max_{x\in\mathcal{X}} v\cdot x - t_{\mathcal{X}}(x)$ also solve the problem $\max_{x\in\bar{\mathcal{X}}} v\cdot x - t(x)$.

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

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

and substitute $x = G(b)$. Similarly, in the hedonic demand model, if the price function, $t(x)$, were non-convex, then certain choices would be sub-optimal under the linear random co-efficients structure.

4.2 Characterization of the Identified Set

The next result characterizes the identified set of type distributions $F_{V|Z}$ given data on the conditional distribution of X given Z. This result does not impose any further restrictions, regularity or otherwise, on either the set $\mathcal{X}$ or $t(\cdot)$. For example, the choice set may be discrete implying that $\mathcal{X}$ is not 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 either differentiable or not.

Stating this result requires some additional notation. Let $\mathcal{V}^{\dagger} \subseteq \mathbb{R}^{J}$ be the open set of points where the conjugate of $t(x)$ given by $t^{*}(v) = \max_{x \in \mathcal{X}} v \cdot x - t_{\mathcal{X}}(x)$ is differentiable, and let $\mathcal{X}^{\dagger} \subseteq \bar{\mathcal{X}}$ as the collection of gradients of t^{*} over $\mathcal{V}^{\dagger}$, i.e. $\mathcal{X}^{\dagger} = \left\{x \in \bar{\mathcal{X}} : x \in \partial t^{*}(v), v \in \mathcal{V}^{\dagger}\right\}$. Observe that if $v \in \mathcal{V}^{\dagger}$, then $x \in \partial t^{*}(v)$ is the unique solution to the problem $\max_{x \in \bar{\mathcal{X}}} v \cdot x - t(x)$. Moreover, if $x \in \mathcal{X}^{\dagger}$, then there exist $v \in \mathcal{V}^{\dagger}$ such that x is the unique solution to the problem $\max_{x \in \bar{\mathcal{X}}} v \cdot x - t(x)$. Thus, $\mathcal{V}^{\dagger}$ is the set of values with a strictly preferred outcome and $\mathcal{X}^{\dagger}$ is the set of outcomes that are strictly preferred by some $v \in \mathcal{V}^{\dagger}$.

Proposition 2. *Suppose Assumption 1 is satisfied. (i) The sharp identified set $\mathcal{F}_{I}(H) = \mathcal{R}_{I} = \mathcal{O}_{I}$ where*

$$\mathcal{R}_{I} = \left\{F_{V,Z} : \textit{ for all } z \textit{ and } \mathcal{X}' \subseteq \mathcal{X}, \Pr\left(X \in \mathcal{X}'|z\right) \leq \int 1\left\{v \in \cup_{x \in \mathcal{X}'} \partial t(x)\right\} \mathrm{d}F_{V|Z=z}\right\},$$

and, if $\mathcal{B}(\mathbb{R}^{n})$ denotes the standard Borel space on $\mathbb{R}^{n}$,

$$\mathcal{O}_{I} = \left\{F_{V,Z} : \textit{ for all } z \textit{ and } \mathcal{V} \in \mathcal{B}(\mathbb{R}^{n}), \int_{\mathcal{V}} \mathrm{d}F_{V|Z=z} \leq Pr\left(X \in \cup_{v \in \mathcal{V}} \partial t^{*}(v)\,|z\right)\right\}.$$

(ii) If $F_{V|Z=z} \in \mathcal{F}_{I}(H)$ admits a density $f_{V|Z=z}$ then $P\left(X \in \mathcal{X}^{\dagger}|z\right) = 1$ and $\Pr\left(X \in \mathcal{X}'|z\right) = \int 1\left\{v \in \cup_{x \in \mathcal{X}'} \partial t(x)\right\} \mathrm{d}F_{V|Z=z}$ for all z and $\mathcal{X}' \subseteq \mathcal{X}$. (iii) If $P\left(X \in \mathcal{X}^{\dagger}|z\right) = 1$ then $\mathcal{F}_{I}$ is not empty.

Proof. Part (i): The conditions test implications of Assumption 1. Thus, $\mathcal{F}_{I}(H) \subseteq \mathcal{R}_{I} \subseteq \mathcal{O}_{I}$. We need to show that $\mathcal{O}_{I} \subseteq \mathcal{F}_{I}(H)$. Consider $F \notin \mathcal{F}_{I}(H)$ and $K_{X^{*}|V,Z} \in \mathbb{K}$. There exist z and $\mathcal{X}' \subseteq \mathcal{X}$ such that $\Pr\left(X \in \mathcal{X}'|z\right) > \int\int 1\left\{X^{*} \in \mathcal{X}'\right\} \mathrm{d}K_{X^{*}|V,z} \mathrm{d}F_{V|Z=z}$ holds. Let $\mathcal{V} = \left\{\cup_{x \in \mathcal{X}'} \partial t(x)\right\}^{c}$, and consider $x \in \cup_{v \in \mathcal{V}} \partial t^{*}(v)$. There exist $v \in \mathcal{V}$ such that $v = \partial t(x)$. Towards a contradiction, assume that $x \in \mathcal{X}'$. Thus, $v \in \cup_{x \in \mathcal{X}'} \partial t(x)$; contradicting $v \in \mathcal{V}$. Thus, $x \in \cup_{v \in \mathcal{V}} \partial t^{*}(v)$ implies $x \notin \mathcal{X}'$. It follows that

$$\Pr\left(X \in \cup_{v \in \mathcal{V}} \partial t^{*}(v)\,|z\right) \leq \Pr\left(X \notin \mathcal{X}'|z\right) < \int\int 1\left\{X^{*} \notin \mathcal{X}'\right\} \mathrm{d}K_{X^{*}|V,z} \mathrm{d}F_{V|Z=z} \leq \int_{\mathcal{V}} \mathrm{d}F_{V|Z=z},$$

where the first inequality follows from set inclusion, the second one is due to the inequality above, and the third one follows because the support of $X^{*}|V = v, z$ is $\partial t^{*}(v)$. Thus, if the inner integral is positive, there exist $x \notin \mathcal{X}'$, such that $x \in \partial t^{*}(v)$.

Part (ii): When $F_{V|Z=z}$ admits a density, the complement of $\mathcal{V}^{\dagger}$ relative to $\mathbb{R}^{J}$ is a set of measure zero Rockafellar (1970, Theorem 25.5). Thus, ties occur with zero probability.

$$\Pr\left(X \in \mathcal{X}|z\right) \leq \int 1\left\{v \in \cup_{x\in\mathcal{X}'}\partial t\left(x\right)\right\} f_{V|Z=z}\left(v\right)\mathrm{d}v = \int 1\left\{v \in \cup_{x\in\mathcal{X}'}\partial t\left(x\right)\cap\mathcal{V}^{\dagger}\right\} f_{V|Z=z}\left(v\right)\mathrm{d}v \leq \Pr\left(X \in\right.$$

The first inequality follows from Proposition 1. The first equality follows because $\Pr\left(V \in \mathcal{V}^{\dagger}|z\right) = 1$. The second inequality follows from the fact that $v \in \partial t\left(x\right)\cap\mathcal{V}^{\dagger}$ implies that x is the unique solution to $\max_{x\in\bar{\mathcal{X}}} v\cdot x - t\left(x\right)$. The last equality follows because $\Pr\left(V \in \mathcal{V}^{\dagger}|z\right) = 1$ implies that $\Pr\left(X \in \mathcal{X}^{\dagger}|z\right) = 1$.

Part (iii). The "only if" part of (ii) is immediate from the arguments showing that $\Pr\left(X \in \mathcal{X}^{\dagger}|z\right) = 1$. The "if" part follows because $\partial t\left(\cdot\right)$ restricted to $\mathcal{X}^{\dagger}$ is invertible. □

Part (i) provides bounds on the generalized likelihood function . The main reason for incompleteness in the model is because it does not specify how ties between different choices in $\mathcal{X}$ are broken. Under the additional restriction that $F_{V|Z}$ admits a density, these ties are measure zero events. The bounds are sharp in this case, and the model delivers a likelihood function. Moreover, as shown in part (iii), the restriction that $F_{V|Z}$ admits a density is falsifiable. Nonetheless, the identified set may 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 points. Computing the subgradients may first require estimating the function $t\left(x\right)$ in some contexts. The second step estimates bounds on the CDF $F_{V|Z}$. 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 a subset of deviations generates a weakly larger identified set. This can be seen by observing that the inequalities $\Pr\left(X \in \mathcal{X}'|z\right) \leq \int 1\left\{v \in \cup_{x\in\mathcal{X}'}\partial t\left(x\right)\right\}\mathrm{d}F_{V|Z=z}$ are imposed for fewer subsets $\mathcal{X}'$ of $\mathcal{X}$. This implication is perhaps expected because fewer necessary conditions for optimality are imposed. Perhaps a less obvious implication is that one can still derive a likelihood function when $F_{V|Z}$ admits a density even if we impose only a subset of deviations. This implication will be important in our analysis of a combinatorial auctions.

A one-dimensional version of part (i) is given in Larsen and Zhang (2018). Specifically, consider an agent, indexed by i, who picks actions that result in a probability of trade equal

[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.

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

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

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

Our result extends naturally to a higher-dimensional version of this problem. It also applies to other scenarios. 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.

[Not sure what to do with this] In higher dimensions, when the support of actions, and consequently the support of x is discrete, the set $\partial t\left(x\right)$ at some of these points in the support has positive volume in R^J. However, if the action space includes both discrete and continuous elements, the set $\partial t\left(x\right)$ may neither be a singleton nor have positive volume in R^J. Instead, it forms a convex set with an affine hull $\operatorname{aff}\left(\partial t\left(x\right)\right)$, that has a dimension $\dim\left(\operatorname{aff}\left(\partial t\left(x\right)\right)\right)$ lying between 0 and J.

4.3 Identification with Rich Choice Environments

Our next result considers the special case when $t\left(x\right)$ is differentiable. It is a corollary of Proposition 1, applied for each fixed value of $Z=z$.

Corollary 1. *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 $t\left(\cdot\right)$ is differentiable for all x in the support of X, then $F_{V,Z}$ is identified.*

Choosing $x\in\mathcal{X}$ uniquely determines the type of all agents making that choice if $t\left(\cdot\right)$ is differentiable at x. For this reason, we label such a choice environment as "rich."

Two special cases of this result are identification of the hedonic choice models and first-price auctions. To see this, recall our calculation of the derivative of $t\left(x\right)=xG^{-1}\left(x\right)$ in first-price auctions (example 3). This derivative, calculated at $x=G\left(b\right)$, is precisely the virtual valuation of a bidder bidding b.

In the higher-dimensional case of hedonic demand models, the marginal cost of increasing x_j is $\partial t\left(x\right)/\partial x_j$. 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 non-linear 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 non-empty interior.*

This assumption allows to consider choice sets that while not being rich in the sense defined above, still provide information on each dimension of payoffs. In particular , if the set $\mathcal{X}$ contains a finite number of options, the function $t(\cdot)$ will be differentiable everywhere except in the set $\mathcal{X}^\dagger \subseteq \mathcal{X}$. However, if the distribution of valuations admits a density, X belongs to $\mathcal{X}^\dagger$ with probability one, which implies that all choices with positive probability result in a non-singleton $\partial t(x)$. Assumption 3 still requires some limited richness of the choice set. A non-empty interior implies that the dimension of $\mathcal{X}$ and of $t(\cdot)$ is J. Thus, the choice of x is informative about every dimension of payoffs.

With these assumptions, we have the following result:

Theorem 1. *Suppose that $v = u + z$ and Assumptions 1, 2 and 3 are satisfied. If z has full support on $\mathbb{R}^J$, then $f_U(\cdot)$ is identified. Moreover, $f_U(\cdot)$ is identified if there exist a linear subspace $L \subseteq \mathbb{R}^J$ such that (i) for every x in the support of X, $t(x)$ is differentiable in every direction $d \in L$ and (ii) z has full support on $L^\perp$, the orthogonal complement of L.*

[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.

The condition of the theorem requires that z has sufficient variation in order to trace the tails of the distribution of u. This condition is similar to those used in other arguments that rely on special regressors.

Proof. Lemma 2 implies that there exists a set $A \in \bar{\mathcal{X}}$ that 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. Fix $k > 0$ such that $f_U \in \mathcal{F}_k$ for the true f_U. With a slight abuse of notation, define the operator $A : \mathcal{F}_k \to \mathbb{L}^\infty\left(\mathbb{R}^d\right)$ as

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

Note that $A\left[f_U\right](z) = \Pr\left(X \in A|z\right)$ for the true value of f_U. We will show that $A\left[f\right] = 0$ a.e. implies that $f = 0$ a.e. if $f \in \mathcal{F}_k$. This statement implies that f_U is identified because if $A\left[f_U\right](z) = \Pr\left(X \in A|z\right)$ for two candidate functions f_U, then linearity of the map A implies that the two functions have to be identical.

Towards a contradiction, suppose $A[f] = 0$ and f is nonzero on a set with positive Lebesgue measure. It is without loss to assume that $\cup_{x\in A}\partial t\left(x\right)$ is contained in a simplicial cone because if it is contained in $C + \{z_0\}$ where C is a simplicial cone, then we can re-define the operator above by replacing the argument of $A\left[f\right]$ with $z' = z - z_0$.. Let $\mathcal{N}\left(C\right)$ be the normal cone to C. By the definition of $\mathcal{F}_k$, there exists $\lambda \in \operatorname{int}\mathcal{N}\left(C\right)$ sufficiently small so that $\exp\left(2\pi u \cdot \lambda\right) f\left(u\right)$ is integrable for all $f \in \mathcal{F}_k$. Fix one such value of λ. Rewrite

$$\begin{aligned}
A\left[f\right](z) &= \int 1\left\{u + z \in \cup_{x\in A}\partial t\left(x\right)\right\} f\left(u\right) \mathrm{d}u \\
&= \int 1\left\{u \in \cup_{x\in A}\partial t\left(x\right)\right\} f\left(u - z\right) \mathrm{d}u \\
&= \exp\left(2\pi z \cdot \lambda\right) \int 1\left\{u \in \cup_{x\in A}\partial t\left(x\right)\right\} \exp\left(-2\pi u \cdot \lambda\right) \exp\left(2\pi\left(u - z\right) \cdot \lambda\right) f\left(u - z\right) \mathrm{d}u.
\end{aligned}$$

Since $\exp\left(2\pi z \cdot \lambda\right) > 0$ a.e., $A\left[f\right] = 0$ a.e. $\iff$ $\hat{\chi}_{\partial t(x),\lambda}\left(\xi\right) \cdot \bar{\hat{f}}_\lambda\left(\xi\right) = 0$, where $\bar{\hat{f}}_\lambda$ is the conjugate of the Fourier transform of $f_\lambda\left(u\right) = \exp\left(2\pi u \cdot \lambda\right) f\left(u\right)$ and $\hat{\chi}_{\partial t(x),\lambda}$ is the Fourier transform of $\chi_{\partial t(x),\lambda} = 1\left\{u \in \partial t\left(x\right)\right\} \exp\left(-2\pi u \cdot \lambda\right)$. Since $\bar{\hat{f}}_\lambda\left(\xi\right)$ is continuous, the set of values ξ where $\hat{f}_\lambda\left(\xi\right) \neq 0$ is open. Further, since $\left\|f_\lambda\right\|_1 > 0$, the support of $\bar{\hat{f}}_\lambda\left(\xi\right)$ is non-empty. Therefore, there is a non-empty open set Ξ, such that $\bar{\hat{f}}_\lambda\left(\xi\right) \neq 0$ for all $\xi \in \Xi$. Because $\hat{\chi}_{\partial t(x),\lambda}\left(\xi\right) \cdot \bar{\hat{f}}_\lambda\left(\xi\right) = 0$, it must be that for all $\xi \in \Xi$, $\hat{\chi}_{\partial t(x),\lambda}\left(\xi\right) = 0$. However, this conclusion contradicts Lemma 1 below, which shows that $\hat{\chi}_{\partial t(x),\lambda}$ is not zero on any open set $\Xi \subseteq \mathbb{R}^J$.

If $t(x)$ is differentiable in every direction $d \in L$ for every $x \in \bar{\mathcal{X}}$ then there is an orthonormal transformation T of outcomes and values so that the resulting $\tilde{t}\,(Tx) = t\,(x)$ is differentiable in the last d arguments, where d is the dimension of the subspace L. 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 $v_2 \in \mathbb{R}^d$. Thus, $f_{U_1}(\cdot)$, the distribution of U_2, which is defined as the last d elements of U, is identified from the observed distribution of the last d elements of $\partial t\,(x) - \{z\}$. 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 $u_2 x_2 - t_2\,(x_2|x_1)$, 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 agents solves $\max_{x_1 \in \mathcal{X}_1} v_1 x_1 - t_1\,(x_1|u_2)$, 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 semi-continuous convex function supported over a non-empty convex and compact set $\bar{\mathcal{X}}_1$ by Assumption 1. $f_{U_1|u_2}(\cdot)$ satisfies part (ii) of Assumption 2 for almost all u_2, and $\mathcal{X}_1$ satisfies assumption 3. Thus, by the first part of this Theorem, $f_{U_1|u_2}(\cdot)$ is identified for almost all u_2. Therefore, the joint distribution of U is identified. □

The proof technique is based on Fourier deconvolution methods. Under Assumption 2, the distribution of v is obtained using a convolution of the distributions of u and z. The choice of a specific set A restricts the set of v to $\cup_{x \in A} \partial t\,(x)$. Therefore, by observing $P\,(X \in A|z)$ for various values of z, we obtain information about the distribution of u because $P\,(X \in A|z) = \int 1\,\{u + z \in \cup_{x \in A} \partial t\,(x)\}\, f(u) \mathrm{d}u$.

The result below states the key technical result in the argument, which is based on finding a vector λ such that $\lambda \cdot v > 0$ for all $v \in \partial t\,(x)$ and then working with the Fourier transform of the function $\chi_{\partial t(x),\lambda}\,(u) = 1\,\{u \in \cup_{x \in A} \partial t\,(x)\} \exp\,(-2\pi u \cdot \lambda)$:

Lemma 1. *Suppose $\partial t\,(x)$ has strictly positive Lebesgue measure and $\partial t\,(x)$ is contained in a (closed) simplicial cone. Then, there exists $\lambda \in int\, \mathcal{N}\,(A)$, with $|\lambda|$ arbitrarily small, such that*

$$\hat{\chi}_{C,\lambda}\,(\xi) = \int 1\,\{u \in C\} \exp\,(-2\pi u \cdot (i\xi + \lambda))\, \mathrm{d}u$$

is not zero on any open set $\Xi \subseteq \mathbb{R}^J$.

The formal proof is presented in Appendix A. The argument has three parts. First, we show that $\hat{\chi}_{C,\lambda}(\xi)$, when viewed as a function with complex domain, is holomorphic in a neighborhood of $\mathbb{R}^J$. This result is obtained by verifying the conditions necessary for differentiating under the integral sign when dealing with functions that have complex domain (Theorem 13.8.6(iii) in DieudonnÃ©, 1976). Second, we apply Theorem 5 in Shabat (1992) which implies that if $\hat{\chi}_{C,\lambda}(\xi)$ is zero on an open subset of $\mathbb{R}^J$ then it is zero everywhere. Third, we observe that this second conclusion contradicts the fact that $\chi_{C,\lambda}(u)$ is, up to scale, a density function.

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 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.

A limitation of Theorem 1 is that it requires a linearly separable regressor z that is independent of u. Our next result applies 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 differentiable 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 non-zero 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 non-zero. 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 in observed in the data, this assumption is falsifiable.[8]

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(ii).

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.*

Proof. Define $v^*(g) = \mathbb{E}\left[\max_{x \in \mathcal{X}} x \cdot (u+g) - t(x) \middle| g\right]$. This expectation exists because

$$\begin{aligned}\mathbb{E}\left[\max_{x \in \mathcal{X}} x \cdot (u+g) - t(x) \middle| g\right] &= \mathbb{E}\left[X^*(u+g) \cdot (u+g) - t(X^*(u+g)) \middle| g\right] \\ &\leq \mathbb{E}\left[\left|X^*(u+g)\right| \cdot |u+g| + \left|t(X^*(u+g))\right| g\right],\end{aligned}$$

which is finite since X^* belongs to compact-valued set, u is independent of g and has finite expectation, and $t(\cdot)$ is a continuous function. Since $X^*(v)$ is measurable, we have that

$$v^*(g) = \max_{X:\mathbb{R}^J \to \mathcal{X}} \int \left[X(u) \cdot (u+g) - t(X(u))\right] f_U(u) \, \mathrm{d}u.$$

The equality follows from setting $X(u) = X^*(u+g) \in \arg\max_{x \in \mathcal{X}} x \cdot (u+g) - t(x)$ to show a weak inequality in one direction, and the definition of $v^*(g)$ for the other. Re-writing, we get that

$$v^*(g) = \max_{X:\mathbb{R}^J \to \mathcal{X}} \left[g \cdot \int X(u) f_U(u) \, \mathrm{d}u + \int \left[X(u) \cdot u - t(X(u))\right] f_U(u) \, \mathrm{d}u\right].$$

Observe that the maximand is linear in g, and therefore equidifferentiable with respect to each g_j. Moreover, the partial derivative of the maximand with respect to each g_j is uniformly bounded because $\mathcal{X}$ is compact. Therefore, by the generalized envelope theorem of Milgrom and Segal (2002) (see Theorem 3),

$$\nabla v^*(g) = \int X^*(u+g) f_U(u) \, \mathrm{d}u = \mathbb{E}\left[X^*(u+g(z)) \,|g(z) = g\right].$$

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

Differentiating, we get that

$$\frac{\partial \mathbb{E}\left[\left.X_k\right|z\right]}{\partial z_l}=\frac{\partial \mathbb{E}\left[\left.X_k^*\left(u+g\left(z\right)\right)\right|z\right]}{\partial z_l}=\partial_{l,k}v^*\left(g\left(z\right)\right)\frac{\partial g_l\left(z_l\right)}{\partial z_l},$$

where the derivatives exist by Assumption 4(iii). Therefore, for any value of g and z, and any pair k and l, we can identify

$$\frac{\partial g_l\left(z_l\right)}{\partial z_l}/\frac{\partial g_k\left(z_k\right)}{\partial z_k}. \tag{3}$$

The rest of the proof uses the arguments in Corollary S.5.1 in Allen and Rehbeck (2017). First, we identify $\frac{\partial}{\partial z_j}g_j\left(z_0\right)$ up to scale. Optimality and independence of u and z implies that

$$\begin{aligned}\left(\mathbb{E}\left[\left.X^*\left(u+g\left(z\right)\right)\right|z\right]-\mathbb{E}\left[\left.X^*\left(u+g\left(z'\right)\right)\right|z'\right]\right)\cdot g\left(z\right)\geq&\left(\mathbb{E}\left[\left.X^*\left(u+g\left(z\right)\right)\right|z\right]\cdot u-\mathbb{E}\left[\left.X^*\left(u+g\left(z'\right)\right)\cdot u\right|z'\right]\right)\\&-\left(\mathbb{E}\left[\left.t\left(X\left(u+g\left(z\right)\right)\right)\right|z\right]-\mathbb{E}\left[\left.t\left(X^*\left(u+g\left(z'\right)\right)\right)\right|z'\right]\right).\end{aligned}$$

An identical expression holds in which the left hand side switches the roles of z and z'. Adding these two inequalities, we get that

$$\left(\mathbb{E}\left[\left.X^*\left(u+g\left(z\right)\right)\right|z\right]-\mathbb{E}\left[\left.X^*\left(u+g\left(z'\right)\right)\right|z'\right]\right)\cdot\left(g\left(z\right)-g\left(z'\right)\right)\geq 0.$$

If $\mathbb{E}\left[\left.X^*\left(u+g\left(z\right)\right)\right|z\right]\neq\mathbb{E}\left[\left.X^*\left(u+g\left(z'\right)\right)\right|z'\right]$, then it must be that $g\left(z\right)\neq g\left(z'\right)$ and

$$\left(\mathbb{E}\left[\left.X^*\left(u+g\left(z\right)\right)\right|z\right]-\mathbb{E}\left[\left.X^*\left(u+g\left(z'\right)\right)\right|z'\right]\right)\cdot\left(g\left(z\right)-g\left(z'\right)\right)>0. \tag{4}$$

Now, consider a small change in the $j-$th component starting from z_0. We know that

$$\left(\mathbb{E}\left[\left.X^*\left(u+g\left(z_0\right)\right)\right|z_0\right]-\mathbb{E}\left[\left.X^*\left(u+g\left(z_0+\Delta_j\right)\right)\right|z_0+\Delta_j\right]\right)\cdot\left(g\left(z_0\right)-g\left(z_0+\Delta_j\right)\right)>0,$$

where Δ_j is the $j-$th standard basis vector multplied by a small $\Delta>0$. Because regressors are dimension-specific, we have that the sign of $g_l\left(z_{0,j}\right)-g_l\left(z_{0,j}+\Delta\right)$ is identified for all $\Delta>0$. 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 (3) 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 $g\left(z\right)$ up to scale and location, assume that there are two functions $g\left(z\right)$ and $\tilde{g}\left(z\right)$ 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. $\square$

The main observation is an envelope theorem argument to show that the gradient of $v^*(g)$ is equal to the expected value of X, which is observed. This argument is akin to Roy's identity from consumer theory. The second derivatives then can be identified and can be used to learn about the derivatives of g.

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.[9]

Theorems 1 and 3 immediately imply the following result:

Corollary 2. *Suppose the hypotheses of Theorem 1 and 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. We describe data and relevant institutional details in section 5.1. In section 5.2, this auction will be modeled as a special case of example 4, Mechanisms with Private Information and Independent Types and Quasi-linear Utilities. Thus, the theoretical results in section 4 will be directly applicable. Section 5.3 then develops an estimation procedure that parallels our identification analysis. Finally, section 5.4 presents our empirical results.

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

[9] We found this direct proof to be simpler than applying the general conditions of Theorems 1 and 2 in 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
# TUs allocated	32
# bidders	20
# bids submitted	43,136
# incumbents participating	8
# TUs with incumbent participating	30
# TUs with incumbent participating + bidding on	30
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.

to schools in the 90 territorial units—henceforth "units". The typical contract has a three-year duration and covers one or multiple units. Epstein et al. (2002) and Kim et al. (2014) provide additional institutional details; we continue to abbreviate the latter as KOW.

The auctioneer uses an annual combinatorial auction to procure meals for about a third of the units. A bid consists of a list of units, i.e., a package, and a per-meal quote. Bidders have to meet some technical and financial requirements to qualify for the auction. Depending on their qualifications, bidders may face restrictions on the packages they can bid, such as a maximum number of units or a maximum number of meals. The number of bids a single bidder can submit is limited to 10,000. The auctioneer determines the contract allocation by solving a linear program that minimizes the total cost and ensures contracts for all units.

We use the same dataset as in KOW. The dataset contains all the bids submitted to the auctions, unit-specific geographical 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 over 43,000 packages. Nine bidders won at least one unit. The average bid was 423 pesos per meal or 71 U.S. cents per meal.[10]

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

Table 2: Units, Bidders, and Bids

	Min	25th percentile	Median	75th percentile	Max
			Panel A: TUs		
Meals (millions)	1.64	2.07	2.40	3.09	3.79
# bidders on TU	20	20	20	20	20
# bids on packages containing TU	1,008	4,219	7,204	9,370	11,198
			Panel B: Bidders		
Max meals allowed (millions)	25.92	35.11	37.77	38.77	50.02
Max TUs allowed	1	3	4	8	8
# TUs 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	31,353	31,353
			Panel C: Bids		
# TUs in package	1	4	5	7	8
Price (pesos per meal)	299.59	401.92	421.41	440.35	690.65
Price if incumbent on package	351.90	405.24	421.96	438.62	608.61

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.

Table 2 describes the bids in greater detail. Panel A shows that there is significant heterogeneity in the total number of meals across units, with the 25th percentile unit serving just over 2 million meals and the 75th percentile serving over 3 million meals. The number of package bids containing a unit also varies substantially. Each unit receives at least one bid from each bidder. Panel B shows that bidders face varying restrictions on the maximum number of meals they are allowed to bid and the maximum number of unit they can be allocated. The number of packages on which a bidder does place a bid is much lower than the number of feasible packages. Finally, panel C shows that the bid price per meal exhibits significant heterogeneity across bidder-package pairs.

5.2 Empirical Model

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. For each bidder i, the outcome space $\mathcal{X}$ consists of vectors $x_i = \{x_{ij}\} \subseteq [0,1]^J$, where x_{ij} denotes bidder i's probability of being allocated package j. Bidder i's action $a_i \in \mathcal{A}_i \subseteq [\mathbb{R}_+ \cup \infty]^J$ specified a bid for each package. A bidder may choose to bid on a subset of packages. We adopt the convention that the bid on a package is infinite if the bidder does not bid on that package.

The Chilean 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 the amount of their winning bids: $\tilde{t}_i (a_i, a_{-i}) = -a_i \cdot \tilde{x}_i (a_i, a_{-i})$. The chosen allocation is determined by finding the lowest total price at which all units can be served. This allocation can be found by solving an integer programming program:

$$\tilde{x} (a_1, \ldots, a_I) \in \underset{x}{\operatorname{argmin}} \sum_i \sum_j x_{ij} a_{ij} \qquad (5)$$
$$\text{s.t. } x \in \tilde{\mathcal{X}},$$

where $x = \{x_{ij}\}_{i,j}$ and $\tilde{\mathcal{X}} \subseteq \mathcal{X}^I$ is a set of allowable allocations. The constraints on $\tilde{\mathcal{X}}$ are (i) each unit k must be served – for each k, $\sum_i \sum_{j:k \in j} x_{ij} \geq 1$; (ii) no bidder wins more than one package – for each i, $\sum_j x_{ij} \leq 1$; and (iii) a set of constraints on each bidder indicating limits on their market share.[11] These constraints, including the third set, are linear in x_{ij}. Ties in the problem are non-generic.

[11]The full set of feasibility constraints is presented in appendix B.1.

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:

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—$c_i \in \mathbb{R}^{2^K-1}$. The empirical implementation in section 5.3 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 $\sigma(c_i; z_i, z_{-i}) : \mathbb{R}^L \to \Delta\mathcal{A}_i$ 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

$$x_{\mathcal{A}}(a_i) = \int \tilde{x}_i(a_i, a_{-i})\, \mathrm{d}F_{A_{-i}|z_i,z_{-i}}(a_{-i}) \tag{6}$$

and (pay-as-bid) expected payment function

$$t_{\mathcal{A}}(a_i) = -a_i \cdot x_{\mathcal{A}}(a_i), \tag{7}$$

where $F_{A_{-i}|z_i,z_{-i}}(a_{-i}) = \prod_{k\neq i} F_{A_k|z_i,z_{-i}}(a_k)$ is the distribution of bidder i's opponents' bids conditional on observables, where the equality follows in a type-symmetric equilibrium from assumption 5. Equilibrium implies that for each bidder i and bid vector a_i submitted by a bidder,

$$\begin{aligned} a_i &\in \operatorname*{argmax}_{a\in\mathcal{A}} -c_i \cdot x_{\mathcal{A}}(a_i) - t_{\mathcal{A}}(a_i) \\ &= \operatorname*{argmax}_{a\in\mathcal{A}} (-c_i + a) \cdot x_{\mathcal{A}}(a) \end{aligned} \tag{8}$$

Thus, the model fits example 4 in section 3.

5.3 Estimation Procedure

Our objective is to estimate the distribution of types $F_{C|z_i}$ given each bidder i's observables z_i. We will use a three-step procedure that parallels our identification arguments. Specifically, in step 1, we will show how to estimate $x_{\mathcal{A}}(a)$ and $t_{\mathcal{A}}(a)$ for any a; in step 2, we will compute revealed preference bounds by calculating $\partial t(x)$ at observed choices $x = x_{\mathcal{A}}(a)$; and in step 3, we will estimate the type distribution $F_{C|z_i}$.

5.3.1 Step 1: Calculating $x_{\mathcal{A}}(a)$ and $t_{\mathcal{A}}(a)$

As shown in equation (6), $x_{\mathcal{A}}(a)$ can be calculated by solving the problem in equation (5) if we have estimates of the distribution of opponent bids $F_{A_{-i}|z_i,z_{-i}}(a_{-i}) = \prod_{k\neq i} F_{A_k|z_i,z_{-i}}(a_k)$. Without further restrictions, the bids A_k are high-dimensional because the number of components equals the number of packages. We follow KOW when modeling and estimating the distribution $F_{A_k|z_i,z_{-i}}(a_k)$. The main restrictions are that (i) each bidder knows the (exogenous) subset of packages on which their competitors will bid, and (ii) competitors' beliefs about the bid distribution can be specified using a low-dimensional parametric distribution. However, an important objective is to account for the effect of potential complementarities and substitutability in costs in the submitted bids.

Bidder i's per-meal bid on package j (from the perspective of i's competitors) is specified as follows:

$$\frac{a_{ij}}{q_j} = \sum_{k\in j} \frac{q_k}{q_j}\tilde{\beta}^{\mathrm{U}}_{ik} + \sum_{l=1}^{L^{\mathrm{volume}}} \mathbf{1}\{q_j \in \mathcal{Q}_l\}\,\beta^{\mathrm{volume}}_{l,\mathrm{size}_i} + \sum_{l=1}^{L^{\mathrm{density}}} \mathbf{1}\{d_j \in \mathcal{D}_l\}\,\beta^{\mathrm{density}}_{l,\mathrm{size}_i} + \epsilon_{ij}, \tag{9}$$

where size_i is a bidder-specific categorical variable taking values "small" and "large." The first term on the right-hand side is a volume-weighted sum of **base prices** $\tilde{\beta}^{\mathrm{TU}}_{ik}$ charged for each constituent unit k. The base prices are distributed as

$$\tilde{\beta}^{\mathrm{Unit}}_{ik} = \beta^{\mathrm{Unit}}_{k} + \beta^{\mathrm{incumb}}\mathrm{incumb}_{ik} + \omega_{ik}, \tag{10}$$

where incumb_{ik} is an indicator for whether bidder i is the incumbent supplier for unit k.[12] The base prices depend on a unit-specific mean $\left(\beta^{\mathrm{Unit}}_{k}\right)_k$, an incumbency shifter β^{incumb}, and a normally distributed random variable $\omega_i = (\omega_{i1}, \ldots, \omega_{iK}) \sim N(0, \Sigma_\omega)$ that allows for correlation across units within a bidder. The next two terms on the right-hand side are **bid adjustments** that capture the effects of cost complementarities or economies of scale and scope due to volume and density. A package's **volume** $q_j = \sum_{k\in j} q_k$ is the annual number of meals to be supplied (in millions), where q_k is the volume of unit k. We specify **density** using

[12] The third set of allocation constraints in the auction limits how many units each bidder is allowed to win. No bidder is allowed to win more than eight units. We call a bidder small if it is allowed to win at most six units and large if it is allowed to win more that six.

a measure of geographic concentration of units within a package that is equivalent to the Herschman-Herfindahl Index. The measure takes values in the unit interval.[13] We discretize volume and density into bins $\{\mathcal{Q}_0, \mathcal{Q}_1 \ldots, \mathcal{Q}_{L^{\text{volume}}}\}$ and $\{\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_{L^{\text{density}}}\}$, respectively, and omit the smallest volume bin and largest density bin during estimation. The bid adjustment amounts $\beta^{\text{volume}} = \left\{\beta^{\text{volume}}_{l,\text{size}}\right\}_{l,\text{size}}$ and $\beta^{\text{density}} = \left\{\beta^{\text{density}}_{l,\text{size}}\right\}_{l,\text{size}}$ depend on bidder size. The final term is an idiosyncratic error ϵ_{ij}; it is normally distributed with mean zero and variance $\sigma^2_{\epsilon,|j|}$ that depends on the number of units in the package.[14]

We estimate the parameters of the bid distribution $\theta_a \equiv (\theta_{a,1}, \theta_{a,2})$ in two stages. First, we estimate $\theta_{a,1} = \left(\beta^{\text{volume}}, \beta^{\text{density}}, \left(\sigma^2_{\epsilon,n}\right)_n\right)$ and unit base prices $\left(\tilde{\beta}_{ik}\right)_{ik}$ from equation (9) using feasible generalized least squares (FGLS). Second, we estimate $\theta_{a,2} = \left(\left(\beta^{\text{TU}}_k\right)_k, \beta^{\text{incumb}}, \Sigma_\omega\right)$ from equation (10) again using FGLS, taking the first-stage $\left(\tilde{\beta}_{ik}\right)_{ik}$ 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 (6) and (7), respectively. For each bidder i, we use $S_a = 1{,}000$ competitor bids a_{-i} from the distribution $F_{A_{-i}|z_i,z_{-i}}\left(\cdot; \hat{\theta}_a\right)$.

5.3.2 Step 2: Bounding Types by Estimating $\partial t(a)$

Restricting the Dimension of the Type Space: Combinatorial auctions attempt to take advantage of cost complementarities across units. By allowing package bidding, they seek to mitigate the exposure problem present in simultaneous or sequential single-unit auctions.[15] The effectiveness of package bidding depends on the magnitude of and across-bidder heterogeneity in these complementarities, as well as on the degree of pass-through from cost complementarities to bid discounts. However, as with the bids, the type $v_i = -c_i$ is high-dimensional. Moreover, the dimensionality $J = 2^K - 1$ grows exponentially in the number of units, K. This curse of dimensionality creates challenges in numerically calculating the

[13] Specifically, the density of package j as $d_j = \sum_r \left(\frac{\sum_{k\in j\cap r} q_k}{q_j}\right)^2$, where the units have been partitioned into regions indexed by $r \in \{1, \ldots, R\}$. It is equal to the probability that two randomly selected meals from package j will come from the same region.

[14] 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] As an extreme example, suppose a bidder views units A and B as perfect complements: its cost of supplying the two-unit package is 400 pesos per meal, while its cost of supplying either of the single-unit packages is infinite. Without package bidding, the bidder always faces the risk of being allocated only one of the units, a prospect that it finds infinitely undesirable. With package bidding, the bidder can choose to bid on the two-unit package but not the single-unit packages, so that it will never be allocated only a single unit.

revealed preference bounds $\partial t\left(a_i\right)$ when the bidder submits a_i. Even if calculating these bounds were feasible, estimating the distribution of costs $F_{C|z_i}$ would be difficult.

Our approach to reducing the dimension of types while capturing complementarities is motivated by the dimension reduction implicit in the specification of bid distribution in equation 9 above. Specifically, we assume that bidder i's per-meal cost $\frac{c_{ij}}{q_j}$ of supplying package j is

$$\frac{c_{ij}}{q_j}=\sum_{k\in j}\frac{q_k}{q_j}\gamma_{ik}^{\text{Unit}}+\sum_{l=1}^{L^{\text{volume}}}\mathbf{1}\left\{q_j\in\mathcal{Q}_l\right\}\gamma_{il}^{\text{volume}}+\sum_{l=1}^{L^{\text{density}}}\mathbf{1}\left\{d_j\in\mathcal{D}_l\right\}\gamma_{il}^{\text{density}}. \tag{11}$$

The per-meal cost for package j is therefore the sum of (i) the volume-weighted average of unit-specific costs $\gamma_{ik}^{\text{Unit}}$ of supplying unit k in the package, (ii) cost complementarities or substitutabilities due to economies or diseconomies of scale, and (iii) the same due to economies or diseconomies of density. Equation (11) above parallels equation (9), which parameterizes the distribution of bidder i's bid on package j. Bidder i's type is therefore $\gamma_i=\left(\left(\gamma_{il}^{\text{volume}}\right)_l,\left(\gamma_{il}^{\text{density}}\right)_l,\left(\gamma_{ik}^{\text{Unit}}\right)_k\right)\in\mathbb{R}^L$ and bidders' costs $c_i\in\mathbb{R}^J$ are a linear function $c_i=M\gamma_i$ of lower-dimensional types $\gamma_i\in\mathbb{R}^L$ where $L\ll J$.

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

Calculating $\partial t\left(a\right)$: Our estimates of ∂t will utilize the linear structure above, while leaving the joint distribution $F_{\Gamma_i|z_i}$ of γ_i given z_i unspecified. We will fit a parametric model to these bounds in a subsequent step. The optimality condition (8) substituted with $c_i=M\gamma_i$ forms the basis for our estimation strategy. Given that bidder i submits a_i, define the **type bounds** $\mathcal{G}_i\subseteq\mathbb{R}^L$, the set of types γ_i for which $a_i\in\operatorname{argmax}_{a\in\mathcal{A}}\left(a-M\gamma_i\right)\cdot x_{\mathcal{A}}\left(a\right)$. Observe that $-c_i$ is in the subdifferential $\partial t\left(x\right)$ if and only if $\gamma_i\in\mathcal{G}_i$. The type γ_i belongs to $\mathcal{G}_i$ if and only if for all package bids a_i',

$$M\gamma_i\cdot\left[x_{\mathcal{A}}\left(a_i\right)-x_{\mathcal{A}}\left(a_i'\right)\right]\leq a_i\cdot x_{\mathcal{A}}\left(a_i\right)-a_i'\cdot x_{\mathcal{A}}\left(a_i'\right). \tag{12}$$

Thus, for each deviation a_i', the revealed preference inequality (12) specifies a half-space in the domain of types. The type bounds $\mathcal{G}_i$ are the intersection of these half-spaces for all possible deviations, forming a convex polyhedral set.[16]

[16] Our model with this simplified type space can be recast in terms of our original model. Observe that

There are two implementation challenges. First, there are a large number of potential deviations and it is computationally costly to check all of them. Simulating the expected allocation function $x_{\mathcal{A}}(\cdot)$ requires 1000 solutions of a high-dimensional integer programming problem at each alternative bid a_i'[17] The second challenge is that the expected allocation function $x_{\mathcal{A}}(\cdot)$ is estimated with simulation error. Considering larger deviations – to bid vectors farther away from the observed a_i – allows the differences $x_{\mathcal{A}}(a_i) - x_{\mathcal{A}}(a_i')$ and $a_i \cdot x_{\mathcal{A}}(a_i) - a_i' \cdot x_{\mathcal{A}}(a_i')$ in inequality (12) to be estimated more precisely, decreasing the variance but at the cost of a larger than necessary estimate of the set $\mathcal{G}_i$ (Fu, 2006).

To address these challenges, we estimate $\mathcal{G}_i$ by checking the inequality in equation (12) for a subset of deviations $a_i' \in \tilde{A}_i \subset \mathcal{A}_i$. Observe that this results in a larger set of types that contains the true set $\mathcal{G}_i$. We consider two types of deviating bid vectors a_i'. In each deviation, we change the bidder's bid on a single package j and hold constant its bids on other packages, $a_{i,-j}' = a_{i,-j}$. The two types of deviations are:

1. **Downward deviations:** For each package j in a subset $\tilde{\mathcal{J}}_i$ of possible packages, decrease the bid on package j by 50 pesos per meal.

2. **Grid deviations:** For each single-unit package j (including single-unit packages that the bidder does not bid on in the data), deviate to bidding 300, 350, 400, 450, and 500 pesos per meal on j. This grid of per-meal bids roughly spans the range of observed bids.

We impose two additional constraints on the sets $\mathcal{G}_i$: (i) costs are below bids ($M\gamma_i \leq a_i$) and (ii) unit-specific costs are nonnegative ($\gamma_i^{\mathrm{TU}} \geq 0$). These inequalities don't require simulating the allocation function $x_{\mathcal{A}}(\cdot)$, making them easier to compute.

The revealed preference inequalities and the two additional constraints yield a convex polyhedral set. Each downward deviation implies a lower bound on some linear combination of type dimensions because the deviation increases the win probability for that package. Constraint (i) generates upper bounds on linear combinations of type dimensions.[18] The grid deviations

$v_i \cdot x = -M\gamma_i \cdot x = -\gamma_i \cdot \left(M^T x\right)$. 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 $t_{\hat{X}}(\hat{x}) = \inf\left\{t_{\mathcal{A}}(a),\ a \in \mathcal{A} : \left(M^T x_{\mathcal{A}}(a)\right) = \hat{x}\right\}$ where $t_{\mathcal{A}}(a_i) = -a_i \cdot x_{\mathcal{A}}(a_i)$. The bidder's problem therefore maximizes $-\gamma_i \cdot \hat{x} - t_{\hat{X}}(\hat{x})$.

[17] Checking many deviations also results in a complex description of the set $\mathcal{G}_i$. As the number of halfspaces grows, polyhedral operations like checking set membership (i.e., checking whether a given type vector lies in $\mathcal{G}_i$), projection onto subspaces, and elimination of dimensions become more difficult.

[18] In principle, we could also obtain upper bounds from the revealed preference inequalities implied by upward deviations (increasing per-meal bids on packages in $\tilde{\mathcal{J}}_i$). However, computing these requires simulating $x_{\mathcal{A}}(\cdot)$ and the resulting upper bounds are not that informative conditional on the bid upper bounds. As a result, we do not include upward deviations in $\tilde{\mathcal{A}}_i$.

and constraint (ii) ensure that unit-specific costs are bounded even if there are units that bidder i is not observed to have placed a bid on. The downward deviation packages $\tilde{\mathcal{J}}_i$ are chosen to generate independent variation in each of the type dimensions: volume economies, density economies, and unit costs. We describe the construction of this set in appendix B.3.

5.3.3 Step 3: Estimating the Type Distribution

For a subset of results, we will parametrize the conditional 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 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 θ_γ using a Metropolis-Hastings algorithm with a simulated version of the likelihood of the parameter vector the parameters $\theta_\gamma = \left(\mu^{\text{volume}}, \mu^{\text{density}}, \left(\mu^{\text{TU}}_k\right)_k, \mu^{\text{incumb}}, \Sigma_\nu\right)$ given bidder i's type bounds $\mathcal{G}_i$. [19] This likelihood is given by

$$L\left(\theta_\gamma \mid \mathcal{G}_i\right) = \int 1\left\{\gamma_i \in \mathcal{G}_i\right\} \mathrm{d}F_{\Gamma_i|z_i}\left(\gamma_i; \theta_\gamma\right). \tag{13}$$

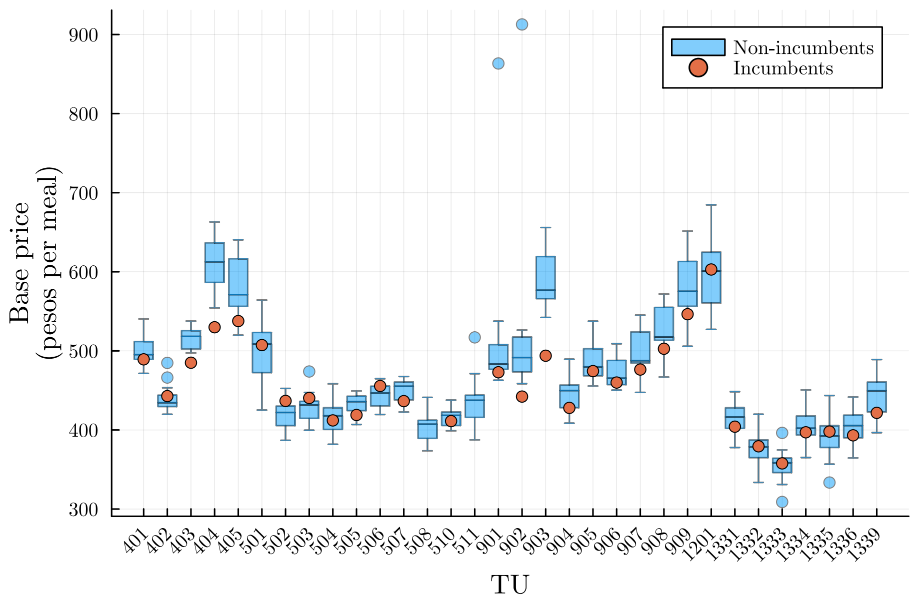
In principle, one can count how many draws from $F_{\Gamma_i|z_i}\left(\cdot \mid \theta_\gamma\right)$ lie in the set $\mathcal{G}_i$, though a precise estimate of a small likelihood—if $F_{\Gamma_i|z_i}\left(\cdot \mid \theta_\gamma\right)$ puts little mass on $\mathcal{G}_i$—may require a very large number of draws. Therefore, we evaluate the likelihood at each candidate θ_γ using importance sampling (Ackerberg, 2009), using draws from a proposal distribution G_i:

$$\begin{aligned} L\left(\theta_\gamma \mid \mathcal{G}_i\right) &= \int 1\left\{x \in \mathcal{G}_i\right\} \frac{f_{\Gamma_i|z_i}\left(x; \theta_\gamma\right)}{g_i\left(x\right)} \mathrm{d}G_i(x) \\ &\approx \frac{1}{S} \sum_{s=1}^{S} 1\left\{x_s \in \mathcal{G}_i\right\} \frac{f_{\Gamma_i|z_i}\left(x_s; \theta_\gamma\right)}{g_i\left(x_s\right)}, \end{aligned} \tag{14}$$

where $\{x_s\}_1^S$ denotes S simulated draws from the proposal distribution G_i. In practice, drawing from the unconditional proposal distribution is computationally burdensome because the high-dimensional nature of this problem implies that obtaining a draw $x_s \in \mathcal{G}_i$ has low

[19] 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. We use a flat prior, so the posterior probability of each candidate θ_γ equals the likelihood. The posterior mode is therefore equivalent to the maximum likelihood estimator.

Figure 2: Distribution of Bidders' Unit Base Prices



Note: The blue boxplots show the distribution of non-incumbent bidders' base prices charged for each unit, denoted β_k^{Unit} in equation (9). 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 charged by the incumbent bidder serving that unit.

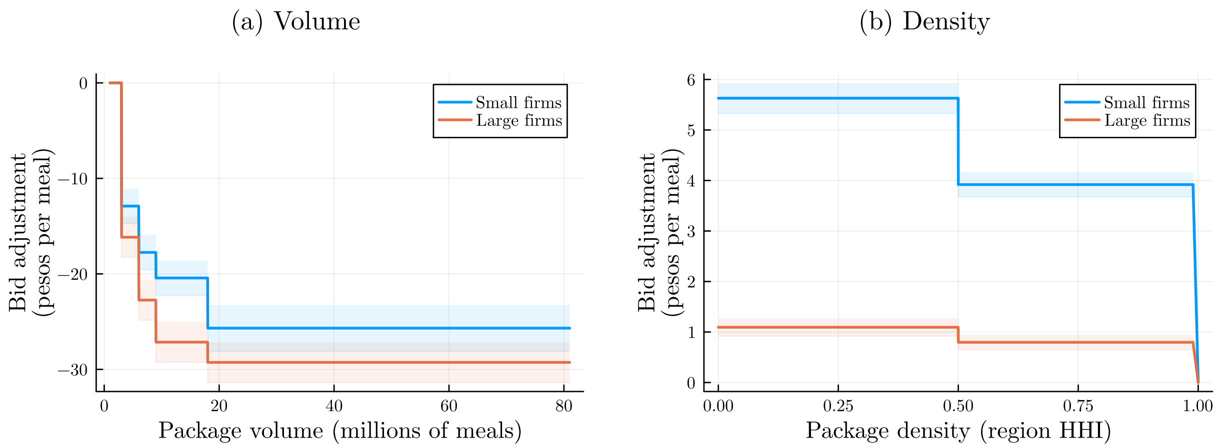
probability. To solve this issue, we take advantage of the fact that the bounds $\mathcal{G}_i$ have been previously computed and draw x_s from a set that is either equal to or slightly larger than $\mathcal{G}_i$. We describe our importance sampling and Metropolis-Hastings procedures in greater detail in appendix B.4.

5.4 Results

5.4.1 Step 1 Estimates

Bid Distribution: Across all bidders and all units, the mean unit-specific base price is 464 pesos per meal, though this masks considerable variation, particularly across units. Figure 2 plots the distribution of unit-specific base prices. The mean base price for unit k, denoted β_k^{Unit} in equation (10), ranges from 356 to 607 pesos per meal. Within a unit, bidders' base prices are correlated. However, we estimate that the incumbent bidder in a unit submits bids

Figure 3: Estimated Bid Adjustment Functions



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 that unit that are 17.62 pesos per meal lower on average than those of its non-incumbent competitors, suggesting a potential incumbency cost advantage.

We find 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, 25.7 to 29.3 pesos lower than those on the smallest packages. Likewise, we estimate that average bids on low-density packages are 1.1 to 5.6 pesos higher than those on high-density ones. Figure 3 shows the estimated bid adjustment functions for small and large bidders. Small bidders offer smaller discounts for package volume than large bidders do. However, the pattern is reversed for density: it is small bidders that offer the greater discounts for package density.

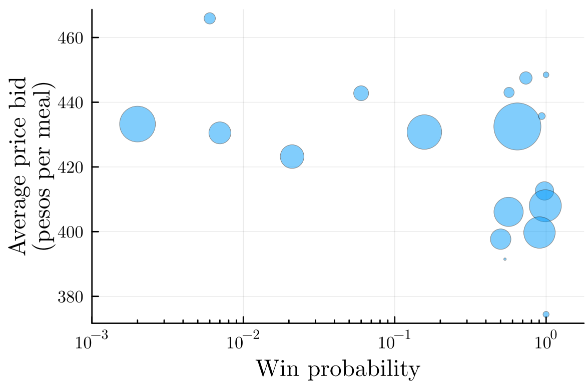
The estimated bid distribution fits the data well; we compare it to the observed bid distribution in figure B.1.

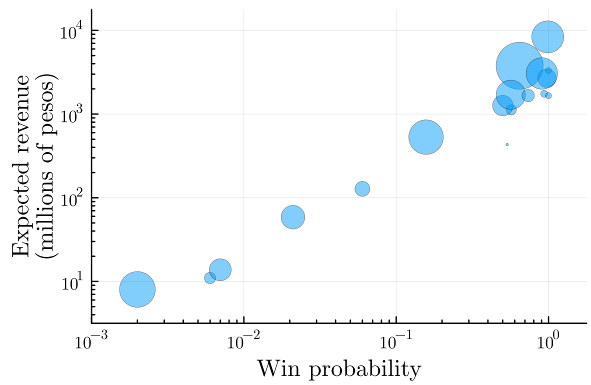
Outcome and Payment Functions: Bidders vary substantially in their probabilities of winning at least one unit. Two bidders are estimated to win with zero probability at their submitted bids, while the win probabilities of another four bidders are estimated to be greater than 95 percent. As expected, bidders' expected revenues $-t_{\mathcal{A}}(a_i)$ increase in their win probabilities, as seen in figure 4. Bidders who submit lower average bids tend to have a

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.

greater probability of winning a package. Although not shown, win probabilities are, perhaps surprisingly, not very correlated with the *number* of bids submitted.

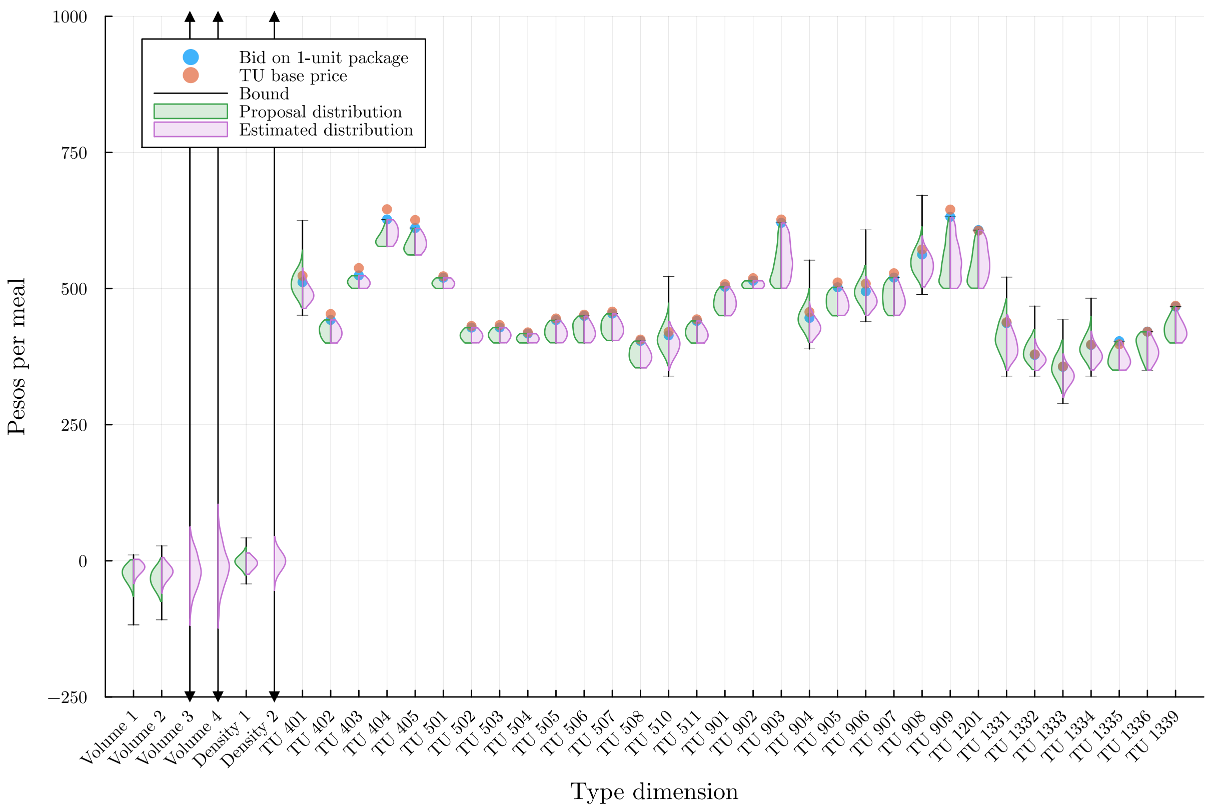
5.4.2 Step 2 Estimates: Type Bounds

Figure 5 shows type bound estimates for an example bidder. There are four volume economies (the lowest volume bin is omitted), two density economies (the highest density bin is omitted), and 32 unit-specific costs. 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 volume bins 3 and 4 or density bin 2, 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.[20] Likewise, this bidder exhibits either small economies or small diseconomies of density. Its unit-specific costs are often but not always bounded from above by both its bid on single-unit packages and the unit-specific base bid.[21]

[20] Higher values for the type dimensions $\gamma_{il}^{\text{volume}}$ and $\gamma_{il}^{\text{density}}$ increase the per-meal package cost in equation (11). 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.

[21] This does not contradict package costs being weakly below package bids because the cost of a single-unit package may also be a function of volume and/or density economies. For example, unit 401 has enough meals to fall in the first (non-omitted) volume bin. If j is the single-unit package containing only unit 401, then

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 type bounds $\mathcal{G}_i$ onto a single dimension. The left green violin plots show the marginals of the proposal distribution G_i. The right purple violin plots show the marginals of the estimated type distribution $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$. Both the proposal distribution and the estimated distribution are conditional on the type bounds $\mathcal{G}_i$.

5.4.3 Step 3 Estimates: Type Distribution

The green violin plots in figure 5 show, for the same example bidder i as above, the marginals of the proposal distribution G_i used to compute the likelihood via importance sampling.[22] The purple violin plots show the marginals of the estimated type distribution $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$. Both distributions are shown conditional on the example bidder's bids – that is, conditional on $\mathcal{G}_i$.[23]

For this example bidder, our estimated type bounds do not rule out economies of volume of up to 118 pesos per meal, but nor do they indicate that such large volume economies (i.e., such negative values for the volume economy dimensions) are likely. Recall that the type bound estimates (black whiskers) plotted in figure 5 are one-dimensional projections of the polyhedral set $\mathcal{G}_i$. Thus, figure 5 shows that there exists a vector of the other type dimensions at which we can rationalize volume economies of 118 pesos per meal. However, the proposal distribution draws place little mass in this region. Since we use a conditional proposal distribution that is uniform on $\mathcal{G}_i$, this indicates that $\mathcal{G}_i$ also has little volume in that region. The estimated type distribution places even less mass in that area; it also has little mass on values of the unit costs above the bidder's per-meal prices bid.[24]

5.5 Markups and Efficiency

5.5.1 Markups in the Combinatorial Auction

Our first set of markup estimates uses only the estimates of the type bounds $\mathcal{G}_i$ and does not require the additional parameterization of the type distribution. For each bid a_{ij} submitted by bidder i on package j, we compute the minimum and maximum markups $\frac{a_{ij}-c_{ij}}{a_{ij}} = \frac{a_{ij}-M_j\gamma_i}{a_{ij}}$ given the bounds on γ_i. The distributions of these lower and upper bounds across bids are shown in blue and orange, respectively, in Figure 6. We bound the aggregate markup on

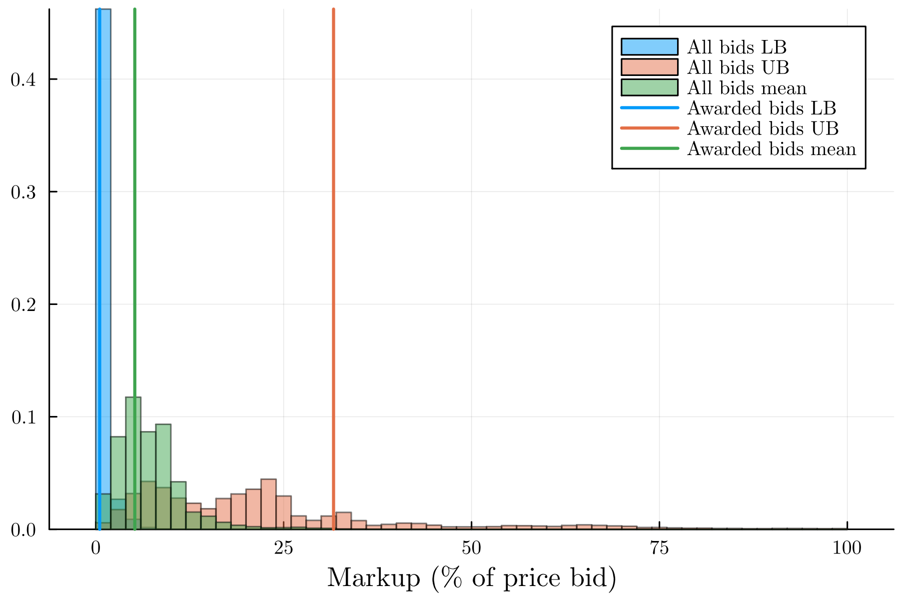
bidder i's per-meal cost of supplying package j is $\frac{c_{ij}}{q_j} = \gamma_{i1}^{\text{volume}} + \gamma_{i,401}^{\text{Unit}} \leq \frac{a_{ij}}{q_{ij}}$. Since for this bidder, we can't rule out negative values of $\gamma_{i1}^{\text{volume}}$, we also can't rule out some unit-specific costs $\gamma_{i,401}^{\text{Unit}}$ above the per-meal price bid $\frac{a_{ij}}{q_j}$.

[22] Tables B.4 and B.5 report summary statistics for the sampled posterior draws for all bidders, and Figure B.2 compares the distributions of bidders' unit-specific base bids $\tilde{\beta}_{ik}$ and unit costs $\gamma_{ik}^{\text{Unit}}$.

[23] We drop this bidder's three infeasible dimensions from the proposal distribution, as we describe in appendix B.4.

[24] We are able to estimate the joint distribution of bidder i's feasible and infeasible type dimensions because estimation of θ_γ pools information from across bidders. However, conditioning on the bounds $\mathcal{G}_i$ is not further informative about the bidder's values of those infeasible dimensions.

Figure 6: Distribution of Estimated Bid Markups



Note: Each histogram observation 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 markups. The green histogram shows the distribution of mean markups. The blue, orange, and green vertical lines respectively indicate the lower bound, upper bound, and mean markups for the awarded allocation.

awarded bids from below by 0.5 percent of the total payment to winning bidders and from above by 31.6 percent.

Second, we compute expected markups 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 $\mathbb{E}\left[\frac{a_{ij}-M_j\gamma_i}{a_{ij}} \mid \gamma_i \in \mathcal{G}_i, z_i; \hat{\theta}_\gamma\right]$. Details on this procedure are provided in appendix B.5. Figure 6 plots the distribution of these expected markups across bidder-package pairs in green. The expected aggregate markup on awarded bids is 5.2 percent. This point estimate is similar to that of KOW, which estimates an aggregate markup of 4.8 percent on awarded bids.

Finally, we report estimates of costs and markups separately for each awarded bid in table B.6.

5.5.2 Efficient Benchmark

Next, we compare the combinatorial auction to a first-best benchmark by simulating counterfactual outcomes under the Vickrey-Clarke-Groves (VCG) mechanism.[25] The VCG allocation $\tilde{x}_i(c_i, c_{-i}) = \left\{x^*_{ij}\right\}_j$ minimizes the total allocation cost, solving

$$x^* \in \underset{x}{\operatorname{argmin}} \sum_i \sum_j x_{ij} c_{ij} \qquad (15)$$
$$\text{s.t. } x \in \mathcal{X}$$

where $x = \{x_{ij}\}_{i,j}$ and $\mathcal{X} \subseteq \prod_i \mathcal{X}_i$ is the same set of allowable allocations imposed in the combinatorial auction (CA). The CA minimizes total payments (the sum of winning bids), while the VCG mechanism minimizes the total social cost, but both are subject to the allocation constraints imposed in Chile. Under the VCG mechanism, each winning bidder is paid the amount of the (positive) externality generated by its participation in the mechanism: $\tilde{t}_i(c_i, c_{-i}) = \left(\min_x \sum_{i' \neq i} \sum_j x_{i'j} c_{i'j}\right) - \left(\sum_{i' \neq i} \sum_j x^*_{i'j} c_{i'j}\right)$. This is the difference between (i) the minimal allocation cost achievable in bidder i's absence and (ii) the minimal allocation allocation cost when i is included.

For computational tractability, we restrict the set of bidder-package pairs that are feasible.[26] A package is considered feasible for a bidder if either (i) at least one bidder bid on that package, (ii) the package has exactly one or two units, (iii) the package is an eight-unit subset of a region with more than eight units, or (iv) the package consists of all units from a single region with no more than eight units. These restrictions imply that the cost of this approximate VCG allocation that we simulate is weakly higher than the unrestricted minimum cost. However, we expect the approximation error to be small because if complementarities amongst other packages not included in this set were substantial, then at least one bidder should have bid the package in the observed auction and the package would be included in group (i).

Not surprisingly, we estimate economically large efficiency gains in moving from the observed CA to the VCG mechanism. The CA allocation cost is estimated to be about 12.2 percent

[25] In the counterfactual analysis, following KOW, we exclude two bidders who submit exceptionally low bids, resulting in high estimated win probabilities and markups. 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 do include these two low-quality bidders in the estimation of the bid and type distributions under the assumption that their competitors did not know they would be disqualified at the time of bidding.

[26] Without any restrictions other than the maximum allowable number of units that can be allocated to a bidder (eight), there are 100,716,104 possible bidder-package pairs.

Table 3: Combinatorial Auction vs. VCG Mechanism

	Mean	SD	p5	p95
	Panel A: Combinatorial auction			
Cost	389.3	2.4	385.4	393.1
Producer surplus	21.6	2.4	17.9	25.6
Payment	411.0	-	-	-
	Panel B: VCG mechanism			
Cost	341.7	8.3	327.3	353.6
Producer surplus	40.4	7.7	29.0	54.3
Payment	382.1	5.7	372.0	390.1
	Panel C: Difference (CA - VCG)			
Cost	47.6	8.1	35.9	61.5
Producer surplus	-18.8	7.6	-32.6	-7.2
Payment	28.9	5.7	20.9	38.9

Note: This table presents summary statistics of simulated CA and VCG outcomes under the estimated type distributions $F_{\gamma_i|z_i}\left(\cdot \mid \hat{\theta}^{MH}\right)$, conditioning on the cost bounds $\mathcal{G}_i$. All values are in pesos per meal.

higher than the (approximate) first-best VCG allocation cost. The VCG mechanism also results in higher producer surplus, defined as payments to bidders less allocation costs. Although total payments to bidders are lower under VCG, this is more than offset by the lower cost of supplying the packages. Table 3 reports summary statistics of simulated allocation costs, producer surpluses, and payments to bidders under the observed combinatorial auction and counterfactual VCG mechanism.

To investigate the source of the efficiency gains, we then recompute allocation costs under each auction design under alternative cost economies, holding fixed the simulated VCG allocation draws. We report the results in table 4. In each panel, the first row shows the same baseline allocation cost as in table 3.[27]

Perhaps the most important source of VCG efficiency gains is the ability to allocate packages with the greatest bidder-specific cost complementarities. Shutting down heterogeneity in

[27] Our estimates of the efficiency gains from the VCG allocation are larger than those obtained by KOW, which finds that the combinatorial auction and the VCG auction yield very similar overall costs. A potential reason for the difference is that KOW directly parametrize markups instead of costs. This approach reduces the heterogeneity in markups relative to our model. If we project our estimated markups on the same set of characteristics used in KOW and volume, and recompute the VCG allocation assuming that only packages that a bidder placed a bid on in the combinatorial auction is considered, then the VCG allocation yields a mean cost of 372.2 pesos per meal, which implies a 4.4 percent VCG efficiency gain. Thus, some of the difference between our estimates and those in KOW that arise from the decision to estimate markups versus directly targeting the distribution of costs.

Table 4: CA and VCG Allocation Costs Under Alternative Cost Economies

	Mean	SD	p5	p95
	Panel A: Combinatorial auction			
Baseline	389.3	2.4	385.4	393.1
Common economies	390.6	3.4	384.9	396.0
Zero economies	406.1	3.4	400.4	411.6
	Panel B: VCG mechanism			
Baseline	341.7	8.3	327.3	353.6
Common economies	371.6	6.0	362.2	381.7
Zero economies	388.1	5.6	379.0	397.8
	Panel C: Difference (CA - VCG)			
Baseline	47.6	8.1	35.9	61.5
Common economies	19.0	6.0	9.0	28.5
Zero economies	18.0	5.7	8.6	27.4

Note: In this table, we fix the simulated VCG allocation draws obtained under the estimated type distributions $F_{\gamma_i|z_i}\left(\cdot \mid \hat{\theta}^{MH}\right)$. In each row in Panels A and B, we compute summary statistics for the CA or VCG allocation cost under alternative cost economies. The baseline case shows the same allocation cost as in table 3. To compute allocation costs under common cost economies, we set the volume and density cost parameters to their across-bidder and across-draw averages. To compute allocation costs under zero cost economies, we set the volume and density cost parameters to be zero. All costs are in pesos per meal.

economies of volume and density barely changes the CA allocation cost but increases the VCG allocation cost by almost 9 percent. The VCG efficiency gain over the CA is reduced from 12.2 percent to 4.9 percent. These alternative allocation costs are computed by setting each bidder's volume and density cost parameters to the across-bidder and across-draw averages of each parameter (see table 4, second row). Eliminating economies of volume and density altogether mechanically increases allocation costs under both mechanisms, but only further reduces the VCG efficiency gain from 4.9 percent to 4.4 percent (see table 4, third row). Thus, the VCG mechanism exploits bidder-specific heterogeneity in economies better than the combinatorial auction.

In theory, through package bidding, these bidder-specific cost complementarities can also show up in the CA. In practice, the pass-through of these complementarities from package costs to package bids is somewhat limited. This occurs in part 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 the multi-product bidder's pricing problem, the 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 presumption for the validity of many statistical procedures (Newey, 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 non-cooperative games; and 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 school choice (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 non-linear multi-dimensional 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 add-on service, which is a common feature of a number of auction settings.

Our results also suggest an estimation approach, which we illustrate by revisiting the combinatorial 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. This provides an alternative approach in a combinatorial auction setting. We estimate that in the 2003 combinatorial auction, the aggregate markup on awarded bids was 5.2 percent. However, a more efficient allocation was possible. We find that the allocation cost under the status-quo combinatorial auction was 12.2 percent higher than it would have been under the first-best Vickrey-Clarke-Groves mechanism. In theory, package bidding in combinatorial auctions allow 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 VCG mechanism's economically large efficiency gains arise from its ability 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 transfer rules out risk aversion. Relaxing this functional form is important. Second, this paper focuses its contributions on identification and largely applies prior methods during estimation. Specifically, we simplified estimation by restricting the dimension of heterogeneity in costs and by parametrizing the cost distribution. Relaxing these assumptions during estimation is also left for future work. 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.

Ackerberg, Daniel A., "A new use of importance sampling to reduce computational burden in simulation estimation," *Quantitative Marketing and Economics*, 12 2009, *7*, 343–376.

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.

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.

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.

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.

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

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

Epstein, Rafael, Lysette HenrÃquez, Jaime CatalÃ¡n, Gabriel Y. Weintraub, and CristiÃ¡n MartÃnez, "A Combinational Auction Improves School Meals in Chile," *Interfaces*, 12 2002, *32*, 1–14.

Fu, Michael C., *Gradient Estimation*, Vol. 13,

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.

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.

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

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.

Milgrom, Paul and Ilya Segal, "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 2002, *70*, 583 – 601.

Newey, Whitney K., "The Asymptotic Variance of Semiparametric Estimators," *Econometrica*, 11 1994, *62*, 1349.

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.

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.

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

Appendix

A Lemmata

Lemma 2. *Let* $t : \mathbb{R}^J \to \mathbb{R} \cup \{\infty\}$ *be a proper closed convex function supported over a convex compact set* $\mathcal{X}$ *that has* has strictly non empty interior *in* $\mathbb{R}^J$. Then there exist $B \subset \mathcal{X}$ such that $\partial t(B) = \left\{v \in \mathbb{R}^J : v \in \partial t(x), x \in B\right\}$ has a non empty interior in $\mathbb{R}^J$ and is contained in a translated simplicial cone.

Proof. The conjugate function t^* is a a proper closed convex function Rockafellar (1970, Theorem 12.2). Because t is proper and lower semi-continuous Rockafellar (1970, page 52), and $\mathcal{X}$is compact, for every $v \in \mathbb{R}^J$, $t^*(v) = \sup_{x \in \mathcal{X}} v \cdot x - t(x)$ is attained for some $x \in \mathcal{X}$. and $\partial t^*(v) \subset \mathcal{X}$. Therefore, the support of t^* is $\mathbb{R}^J$. The set D where the function $f^*$is differentiable is a dense subset of $\mathbb{R}^J$, and the gradient mapping $\nabla t^* : v \to \nabla t^*(v)$ is continuous on D Rockafellar (1970, Theorem 25.5). Take any $v_0 \in D$ and let $x_0 = \nabla t^*(v_0) = \partial t^*(v_0) \in \mathcal{X}$. Let $S = \{x_1, x_2, ..., x_J\} \subset \mathcal{X}$ be such that the convex hull of $S \cup \{x_0\}$ has has strictly positive Lebesgue measure in $\mathbb{R}^J$. Such set S exists because $\mathcal{X}$ has has strictly positive Lebesgue measure in $\mathbb{R}^J$. 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 in $\mathcal{X}$. Let $\bar{B}(x_0)$ be neighborhood of x_0 in the interior of the convex hull of $S \cup \{\bar{x}\}$. By the continuity of the gradient mapping ∇t^*, there exist a neighborhood, $\bar{V}(v_0) \subset \mathbb{R}^J$ such that $B = \left\{x \in \mathbb{R}^J : x = \nabla t^*(v), v \in \bar{V}(v_0)\right\} \subset \bar{B}(x_0)$. Because for all $x \in B$, $x \in \partial t^*(v) \subset \mathcal{X}$ for some $v \in \bar{V}(v_0)$, it follows that $B \subset \mathcal{X}$. Moreover, if $v \in \bar{V}(v_0)$, then $x \in \partial t^*(v) \subset B$ Rockafellar (1970, Theorem 25.6). Therefore, $v \in \partial t(x) \subset \partial t(B)$ Rockafellar (1970, Theorem 23.5). Thus, $\partial t(B) \supset \bar{V}(v_0)$ has has strictly positive Lebesgue measure in $\mathbb{R}^J$. □

The conjugate function t^* is a a proper closed convex function Rockafellar (1970, Theorem 12.2). Because t is proper and lower semi-continuous Rockafellar (1970, page 52), and $\mathcal{X}$is compact, for every $v \in \mathbb{R}^J$, $t^*(v) = \sup_{x \in \mathcal{X}} v \cdot x - t(x)$ is attained for some $x \in \mathcal{X}$. and $\partial t^*(v) \subset \mathcal{X}$. Therefore, the support of t^* is $\mathbb{R}^J$. The set D where the function $f^*$is differentiable is a dense subset of $\mathbb{R}^J$, and the gradient mapping $\nabla t^* : v \to \nabla t^*(v)$ is continuous on D Rockafellar (1970, Theorem 25.5). Take any $v_0 \in D$ and let $x_0 = \nabla t^*(v_0) = \partial t^*(v_0) \in \mathcal{X}$. Let $S = \{x_1, x_2, ..., x_J\} \subset \mathcal{X}$ be such that the convex hull of $S \cup \{x_0\}$ has has strictly positive Lebesgue measure in $\mathbb{R}^J$. Such set S exists because $\mathcal{X}$ has has strictly positive Lebesgue measure in $\mathbb{R}^J$. 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 in $\mathcal{X}$. Let $\bar{B}(x_0)$ be neighborhood of x_0 in the interior of the convex hull of $S \cup \{\bar{x}\}$. By the continuity of the gradient mapping ∇t^*, there exist a neighborhood, $\bar{V}(v_0) \subset \mathbb{R}^J$ such that $B = \left\{x \in \mathbb{R}^J : x = \nabla t^*(v), v \in \bar{V}(v_0)\right\} \subset \bar{B}(x_0)$. Because for all $x \in B$, $x \in \partial t^*(v) \subset \mathcal{X}$ for some $v \in \bar{V}(v_0)$, it follows that $B \subset \mathcal{X}$. Moreover, if $v \in \bar{V}(v_0)$, then $x \in \partial t^*(v) \subset B$ Rockafellar (1970, Theorem 25.6). Therefore, $v \in \partial t(x) \subset \partial t(B)$ Rockafellar (1970, Theorem 23.5). Thus, $\partial t(B) \supset \bar{V}(v_0)$ has has strictly positive Lebesgue measure in $\mathbb{R}^J$.

Now, we show that 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 $\bar{B}(x_0)$ with transfers $t(\cdot)$ would be dominated by the outcomes in $\mathcal{Y}$. 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 with 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. 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 3. *Let C be a subset of $\mathcal{X} \subseteq \mathbb{R}^J$ 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 non-empty because $\bar{C}$ is simplicial. Fix a $\lambda \in \text{int}\, \mathcal{N}\left(\bar{C}\right)$ 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 funtion on $\Xi_\lambda = \{\xi \in \mathbb{C}^J : \xi = y + iz, x \in$

$\mathbb{R}^J, \frac{\lambda}{2} - z \in \operatorname{int} \mathcal{N}(\bar{C})\}$, 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 non-zero because C has strictly positive Lebesgue measure.

The second part follows immediately because if $\lambda \in \operatorname{int} \mathcal{N}\left(\bar{C}\right)$ 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. To do this, we first use the differentiation under the integral sign theorem for complex variables (Theorem 13.8.6(iii) in DieudonnÃ©, 1976) to show that $\frac{\partial \hat{\chi}_{C,\lambda}(\xi)}{\partial \xi_k}$ exists on Ξ_λ and is equal to

$$\frac{\partial \hat{\chi}_{C,\lambda}(\xi)}{\partial \xi_k} = \int 1\{v \in C\} i v_k \exp(-2\pi v \cdot (i\xi + \lambda))\, \mathrm{d}v.$$

Fix $\xi_{-k} = (\xi_1, \ldots, \xi_{k-1}, \xi_{k+1}, \ldots, \xi_J)$ and define the function $f_C(v, \xi_k; \xi_{-k}) = 1\{v \in C\} \exp(-2\pi v \cdot (i\xi + \lambda))$. 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 $\left|f_{\bar{C}+\{x\}}(v, \xi_k; \xi_{-k})\right| \geq |f_C(v, \xi_k; \xi_{-k})|$ for all v and show that $\left|f_{\bar{C}+\{x\}}(v)\right| \leq g(v) = 1\left\{v - x \in \bar{C}\right\} \left|\exp\left(-2\pi v \cdot \frac{\lambda}{2}\right)\right|$ for all $\xi \in \Xi_\lambda$, where $g(v)$ is integrable. To this end, we first bound $|f_{\bar{C}+\{x\}}(v)|$ as follows:

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

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

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

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

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

The first equality uses the definition of $A_{\bar{C}}$. The second substitutes $u = A_{\bar{C}}^{-1}(v - x)$, and uses the fact that $\left|\det\left(A_{\bar{C}}\right)\right| = 1$ and $u \geq 0$. The third rewrites the integrands as the product of two terms and pulls out the constant $\left|\exp\left(-2\pi x \cdot \frac{\lambda}{2}\right)\right|$ from the integral sign.

Since $\exp\left(-2\pi x \cdot \frac{\lambda}{2}\right)$ is finite, we need 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

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

where $\sigma\left(u'\right)$ 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 $\Gamma\left(\cdot\right)$ 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 $h \in \mathbb{C}^J$,

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

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 $1\{v \in C\} \left|\exp\left(-2\pi v \cdot (i(\xi + h) + \lambda)\right) - \exp\left(-2\pi v \cdot (i\xi + \lambda)\right)\right| = \exp(-2\pi v \cdot \lambda) \left|\exp(-2\pi v \cdot ih) - 1\right| \to 0$, as $h \to 0$. Therefore, by the dominated convergence theorem, $|\hat{\chi}_{C,\lambda}(\xi + h) - \hat{\chi}_{C,\lambda}(\xi)| \to 0$ as $h \to 0$. $\square$

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 $x_{ij} \in \{0, 1\}$ indicate whether bidder i is allocated package j. The combinatorial auction minimizes total payments (the sum of winning bids) by solving the integer program

$$
\begin{aligned}
\min_{x} \ & \sum_{i} \sum_{j} x_{ij} a_{ij} \\
\text{s.t.} \ & \sum_{i} \sum_{j:k \in j} x_{ij} \geq 1 \quad \text{for all } k \\
& \sum_{j} x_{ij} \leq 1 \quad \text{for all } i \\
& \sum_{j} x_{ij} q_{ij} \leq \bar{q}_i \quad \text{for all } i \\
& \sum_{j} x_{ij} \left| j \right| \leq \bar{n}_i \quad \text{for all } i \\
& \underline{I}_r \leq \sum_{i} \sum_{j: j \cap r \neq \varnothing} x_{ij} \leq \bar{I}_r \quad \text{for all } r \\
& \sum_{i} \sum_{j} x_{ij} \geq \underline{I}
\end{aligned}
$$

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.

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

	Small firms	Large firms
Volume 1	-12.899	-16.158
	(0.929)	(1.085)
Volume 2	-17.752	-22.742
	(0.934)	(1.077)
Volume 3	-20.429	-27.144
	(0.937)	(1.076)
Volume 4	-25.683	-29.265
	(1.234)	(1.077)
Density 1	3.919	0.797
	(0.125)	(0.076)
Density 2	5.629	1.094
	(0.155)	(0.091)

Note: Standard errors are in parentheses.

B.2 Bid Distribution

B.2.1 Bid Adjustment Functions

The bid adjustment functions $h_a^{\text{volume}}\left(\cdot;\beta_{\text{size}_i}^{\text{volume}}\right)$ and $h_a^{\text{density}}\left(\cdot;\beta_{\text{size}_i}^{\text{density}}\right)$ are step functions in the volume q_j and density d_j, respectively, of the package j. The volume adjustment function is constant in each of nine equally spaced intervals $\{[0,3],\ldots,[24,27]\}$ that cover the range of package volumes observed in the data. The density adjustment function is constant in each of the intervals $\{[0,0.5),[0.5,1),[1,1]\}$, with the third bin capturing the mass of packages whose constitutent units lie entirely within one region.[28] We normalize bid adjustments in the smallest volume bin and largest density bin to zero.

B.2.2 Parameter Estimates

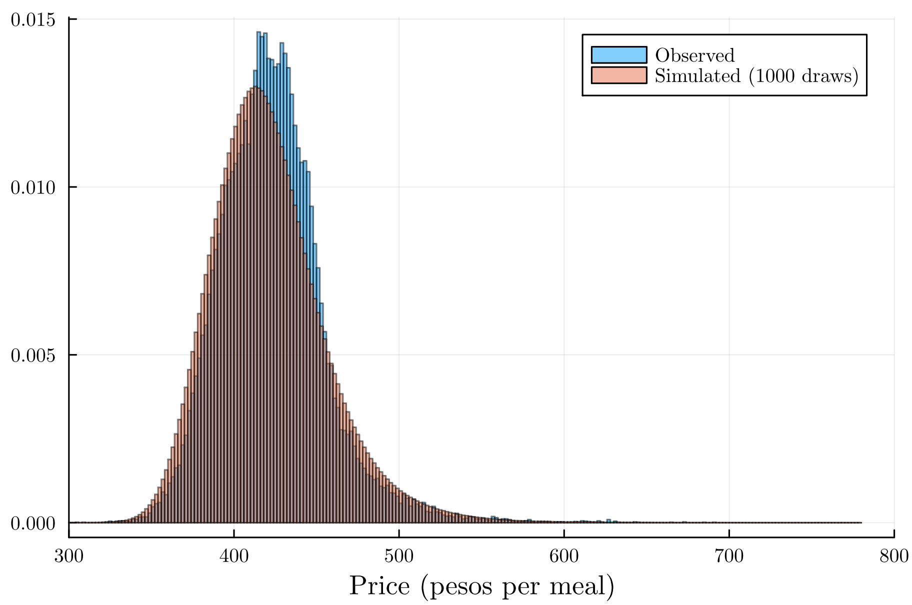
In tables B.1, B.2, and B.3, we present estimates $\hat{\theta}_a$ of the bid distribution parameters. In figure B.1, we compare the observed distribution of per-meal prices bid to our estimated distribution.

[28]Package volume is measured in millions of meals per year. We define package density in footnote 13; it takes values in the unit interval.

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

Package size (# TUs)	Standard deviation
1	13.757
2	4.870
3	3.901
4	3.211
5	2.523
6	2.054
7	1.319
8	1.502

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.

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

	β	σ_ω
TU 401	500.949	14.839
TU 402	440.899	12.154
TU 403	515.466	13.684
TU 404	606.711	33.822
TU 405	580.172	33.026
TU 501	501.379	30.620
TU 502	421.620	19.436
TU 503	430.904	17.976
TU 504	417.353	21.962
TU 505	433.111	11.867
TU 506	445.616	13.667
TU 507	450.197	12.639
TU 508	404.258	14.435
TU 510	416.788	9.913
TU 511	433.948	29.586
TU 901	510.563	25.836
TU 902	512.277	24.509
TU 903	585.829	36.108
TU 904	443.816	18.065
TU 905	486.187	22.731
TU 906	474.471	19.505
TU 907	499.643	24.880
TU 908	528.013	24.865
TU 909	581.423	38.627
TU 1201	603.123	48.260
TU 1331	415.935	13.356
TU 1332	378.521	11.529
TU 1333	356.387	11.532
TU 1334	404.167	11.000
TU 1335	392.525	21.249
TU 1336	404.571	12.959
TU 1339	444.407	20.426
Incumbency shifter	-17.621	

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

B.3 Type Bounds

As part of estimating the type bounds $\tilde{\mathcal{G}}_i$, we rule out downward deviations to bid vectors a_i' in which bidder i decreases its bid on a single package j by 50 pesos per meal. The set of packages $\tilde{\mathcal{J}}_i$ for which we consider these downward deviations is chosen to generate independent variation in each of the type dimensions. Specifically, for each of the type dimensions indexed by l, we start by choosing a set of packages $\tilde{\mathcal{J}}_{il}$ to be maximally informative about that dimension:

- If l is a volume or density economy, then $\tilde{\mathcal{J}}_{il}$ is the set of 10 packages with the highest estimated win probabilities for bidder i under its observed bids.
- If l is a unit cost, then $\tilde{\mathcal{J}}_{il}$ contains only the single-unit package with that unit. If we do not observe bidder i bidding on the single-unit package for that unit, then the sole element of $\tilde{\mathcal{J}}_{il}$ is the next-smallest package i bids on, with ties broken in descending order of i's estimated probability of winning that package.

Then, we take the union of these dimension-specific sets $\tilde{\mathcal{J}}_{il}$ and further restrict to packages that (i) bidder i bids on in the data, (ii) bidder i is allowed to win under the meal and unit constraints described in appendix B.1, and (iii) whose costs have a nonzero coefficient on dimension l.

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.

B.4 Type Distribution

B.4.1 Importance Sampling

In principle, it suffices to estimate the likelihood (13) at each candidate θ_γ by counting how many draws from $F_{\Gamma_i|z_i}(\cdot \mid \theta_\gamma)$ lie in the set $\mathcal{G}_i$, though a precise estimate of a small likelihood – if $F_{\Gamma_i|z_i}(\cdot \mid \theta_\gamma)$ puts little mass on $\mathcal{G}_i$ – may require a large number of draws. We would ideally like to sample γ_i from $F_{\Gamma_i|z_i}(\cdot \mid \theta_\gamma)$ conditional on γ being in $\mathcal{G}_i$, but doing so in closed form is possible only in one dimension. Instead, we use importance sampling, taking draws from a proposal distribution G_i and reweighting them by their likelihood ratios:

$$L(\theta_\gamma \mid \mathcal{G}_i) = \int 1\{x \in \mathcal{G}_i\} \frac{f_{\Gamma_i|z_i}(x;\theta_\gamma)}{g_i(x)} \mathrm{d}G_i(x) \tag{14}$$

We furthermore use draws from G_i conditional on the type bounds $\mathcal{G}_i$; denote this conditional distribution by $\tilde{G}_i$. Inside the set $\mathcal{G}_i$, the unconditional and conditional distributions differ by the total mass that G_i puts on $\mathcal{G}_i$:

$$\mathrm{d}\tilde{G}_i(x) = \frac{1\{x \in \mathcal{G}_i\}}{\int 1\{y \in \mathcal{G}_i\} \mathrm{d}G_i(y)} \mathrm{d}G_i(x) \tag{16}$$

In particular, when $x \in \mathcal{G}_i$. then $\mathrm{d}G_i(x) = \left[\int 1\{y \in \mathcal{G}_i\} \mathrm{d}G_i(y)\right] \mathrm{d}\tilde{G}_i(x)$. As a result, the likelihood in equation (14) becomes

$$\begin{aligned} L(\theta_\gamma \mid \mathcal{G}_i) &= \int 1\{x \in \mathcal{G}_i\} \frac{f_{\Gamma_i|z_i}(x;\theta_\gamma)}{g_i(x)} \left[\int 1\{y \in \mathcal{G}_i\} \mathrm{d}G_i(y)\right] \mathrm{d}\tilde{G}_i(x) \\ &= \left[\int 1\{y \in \mathcal{G}_i\} \mathrm{d}G_i(y)\right] \times \int \frac{f_{\Gamma_i|z_i}(x;\theta_\gamma)}{g_i(x)} \mathrm{d}\tilde{G}_i(x) \\ &=: \left[\int 1\{y \in \mathcal{G}_i\} \mathrm{d}G_i(y)\right] \times \tilde{L}\left(\theta_\gamma \mid \tilde{\Gamma}_i\right) \end{aligned}$$

where the second equality uses the fact that, by definition, $1\{x \in \mathcal{G}_i\} \equiv 1$ for x drawn from $\tilde{G}_i$. Call $\tilde{L}(\theta_\gamma \mid \mathcal{G}_i)$ the **quasi-likelihood**, which differs from the true likelihood by the same factor $\int 1\{y \in \mathcal{G}_i\} g_i(y) \mathrm{d}y$ as in equation (16). Since this factor does not depend on θ_γ, it suffices to use the quasi-likelihood in Metropolis-Hastings rather than the true likelihood. The quasi-likelihood is easy to compute, as it doesn't require integrating out the mass that the unconditional proposal distribution puts on $\mathcal{G}_i$. We approximate the quasi-likelihood using $S_\gamma^{\text{prop}} = 1{,}000$ draws $\gamma_i^{(s)}$ from $\tilde{G}_i$, the version of the proposal distribution that conditions on the type bounds:

$$\tilde{L}\left(\theta_\gamma \mid \tilde{\Gamma}_i\right) \approx \frac{1}{S_\gamma^{\text{prop}}} \sum_{s=1}^{S_\gamma^{\text{prop}}} \frac{f_{\Gamma_i|z_i}\left(\gamma_i^{(s)};\theta_\gamma\right)}{g_i\left(\gamma_i^{(s)}\right)}$$

Appendix B.4.2 describes how we generate draws from the conditional proposal distribution.

Finally, this discussion has so far assumed that bidder i is allowed to win at least one package involving each type dimension. If there are in fact some dimensions that are infeasible for i, then $\mathcal{G}_i$ must always be completely unbounded in those dimensions, as no bids by the bidder can ever be informative about its values for those dimensions. When evaluating the (quasi-)likelihood, we therefore drop infeasible dimensions from both the type bound $\mathcal{G}_i$ and the proposal distribution G_i, and we integrate out over infeasible dimensions in $F_{\Gamma_i|z_i}$. (This is easy because G_i is a product distribution and $F_{\Gamma_i|z_i}$ is multivariate normal.)

B.4.2 Proposal Distribution

Constructing the Unconditional Distribution The advantage of importance sampling is that fewer draws are required to precisely estimate the likelihood; the challenge is choosing the proposal distribution well. Its support must contain $\mathcal{G}_i$, and it ideally concentrates most of its mass there. Its density should be easy to evaluate. The variance of the resulting likelihood ratios should be minimized, so that the effective sample size is maximized. The difficulty of designing the right proposal distribution increases with the dimension of the random vector.

Let G_i be a (bidder-specific) product distribution on the L-dimensional box containing $\mathcal{G}_i$. We use a product of uniform and exponential distributions, depending on whether the given dimension is bounded. For each dimension l, define

$$
\begin{aligned}
\underline{\gamma}_{il} &= \inf\left\{\gamma_{il} \in \mathbb{R} \cup \{-\infty\} : (\gamma_{il}, \gamma_{i,-l}) \in \mathcal{G}_i \text{ for some } \gamma_{i,-l} \in \mathbb{R}^{L-1}\right\} \\
\bar{\gamma}_{il} &= \sup\left\{\gamma_{il} \in \mathbb{R} \cup \{+\infty\} : (\gamma_{il}, \gamma_{i,-l}) \in \mathcal{G}_i \text{ for some } \gamma_{i,-l} \in \mathbb{R}^{L-1}\right\} \\
\underline{\chi}_{il} &= \begin{cases} -100, & l \text{ is a volume or density economy} \\ 300, & l \text{ is a TU cost} \end{cases} \\
\bar{\chi}_{il} &= \begin{cases} 100, & l \text{ is a volume or density economy} \\ 600, & l \text{ is a TU} \end{cases}
\end{aligned}
$$

Then we define G_{il}, the proposal distribution for dimension l, as follows:

- If $\underline{\gamma}_{il} > \underline{\chi}_{il}$ and $\bar{\gamma}_{il} < \bar{\chi}_{il}$, then G_{il} is uniform on $[\underline{\gamma}_{il}, \bar{\gamma}_{il}]$.
- If $\underline{\gamma}_{il} > \underline{\chi}_{il}$ but not $\bar{\gamma}_{il} < \bar{\chi}_{il}$, then we take G_{il} to be unbounded from above. Let G_{il} be the shifted and re-scaled exponential distribution with minimum $\underline{\gamma}_{il}$ and 75th percentile $\bar{\chi}_{il}$.
- If $\bar{\gamma}_{il} < \bar{\chi}_{il}$ but not $\underline{\gamma}_{il} > \underline{\chi}_{il}$, then we take G_{il} to be unbounded from below. Let G_{il} be the shifted, rescaled, and flipped exponential distribution with maximum $\bar{\gamma}_{il}$ and 25th percentile $\underline{\chi}_{il}$.

- If neither $\underline{\gamma}_{il} > \underline{\chi}_{il}$ nor $\bar{\gamma}_{il} < \bar{\chi}_{il}$, then we take G_{il} to be unbounded in both directions. Let G_{il} be the shifted and re-scaled Laplace (double exponential) distribution with 25th percentile $\underline{\chi}_{il}$ and 75th percentile $\bar{\chi}_{il}$.

Ideally, we would like to sample uniformly from the smallest box in $\mathbb{R}^L$ containing $\mathcal{G}_i$, which would amount to sampling uniformly and independently from each dimension. This is certainly not possible in dimension l if $\mathcal{G}_i$ is unbounded in that dimension—that is, if the **outer bounds** $\underline{\gamma}_{il}$ and $\bar{\gamma}_{il}$ are not both finite. In this case, we want that dimension's proposal distribution G_{il} to have full support on the half or full real line, hence we set it to the exponential or double exponential. However, even if $\underline{\gamma}_{il}$ and $\bar{\gamma}_{il}$ are both finite, these outer bounds might still be very uninformative, and it will be undesirable to sample uniformly from such a large interval. If we don't have $\underline{\chi}_{il} < \underline{\gamma}_{il} < \bar{\gamma}_{il} < \bar{\chi}_{il}$, then we still use the exponential or double exponential for G_{il}.

Sampling from the Conditional Distribution We take proposal draws from G_i *conditional on* $\mathcal{G}_i$ using Gibbs sampling to draw one dimension at a time. Suppose we want to sample the sth new value for the lth cost parameter. Let $\gamma_{i,-l}^{(s)} = \left(\gamma_{i1}^{(s)}, \ldots \gamma_{i,l-1}^{(s)}, \gamma_{i,l+1}^{(s-1)}, \ldots, \gamma_{i,n_\gamma}^{(s-1)}\right)$ be the vector of current values for the other $L-1$ dimensions. The set of values γ_{il} such that $\left(\gamma_{il}, \gamma_{i,-l}^{(s)}\right) \in \mathcal{G}_i$ is an interval, so we draw the new cost parameter $\gamma_{il}^{(s)}$from G_{il} truncated to this interval.

We start by sampling 10,000 proposal draws, discarding the first half as burn-in and thinning the remainder by keeping every fifth draw. We end up with $S_\gamma^{\text{prop}} = 1{,}000$ proposal draws which are used to estimate the likelihood.

B.4.3 Metropolis-Hastings

Before running Metropolis-Hastings, we first estimate θ_γ via maximum (quasi-)likelihood. We start the maximum likelihood estimation at the following parameters: $\mu^{\text{volume}} = 0$, $\mu^{\text{density}} = 0$, $\mu^{\text{TU}} = (400, \ldots, 400)$, $\mu^{\text{incumb}} = 0$, and $\Sigma = 50^2 I$. Let $\hat{\theta}_\gamma^{ML}$ denote the MLE point estimate, and let $\hat{H}_\gamma^{ML}$ be the Hessian of the log-(quasi-)likelihood evaluated at $\hat{\theta}_\gamma^{ML}$. We use $\hat{\theta}_\gamma^{ML}$ as the starting point for Metropolis-Hastings.

For the jump distribution, we use a multivariate normal centered at zero, with variance-covariance matrix proportional to the inverse of the negative of the Hessian, $\left(-\hat{H}_\gamma^{ML}\right)^{-1}$. The scaling parameter is chosen to target a rejection rate of between 0.4 and 0.6; in practice, we multiply the inverse Hessian by 0.03^2. The inverse Hessian contains information about how large a step to take in each dimension and how correlated the dimensions are. If the posterior marginal distribution of a parameter is very concentrated, then the corresponding diagonal element of the Hessian will be large, so the step size in that dimension will be small. When the posterior marginal distribution is very diffuse, the step size under the jump distribution will be large.

We start with 1 million Metropolis-Hastings iterations; then, we discard the first half as burn-in and thin the remainder by keeping every fiftieth draw. The result is $S_{\theta_\gamma} = 10{,}000$ draws of θ_γ. We report summary statistics for the sampled draws in tables B.4 and B.5.

Table B.4: Metropolis-Hastings Draws of Volume and Density Economy Parameters

	μ								σ			
	Small firms				Large firms							
			95% confidence set				95% confidence set				95% confidence set	
	Mean	SD	LB	UB	Mean	SD	LB	UB	Mean	SD	LB	UB
Volume 1	-11.507	21.928	-94.290	70.663	-9.275	26.542	-115.293	122.521	74.204	12.680	42.428	131.952
Volume 2	-30.158	11.590	-73.483	21.540	-8.649	14.055	-71.863	47.702	37.765	8.048	20.819	94.444
Volume 3	-19.683	11.886	-66.354	35.007	-19.900	11.456	-64.250	23.391	30.848	7.087	16.193	81.524
Volume 4	-9.298	37.740	-215.879	148.431	-25.790	18.891	-108.725	74.027	35.638	17.707	6.106	129.513
Density 1	-4.600	3.435	-18.132	8.625	2.711	4.137	-13.423	18.375	10.853	2.119	5.646	19.080
Density 2	0.178	6.126	-26.866	41.199	4.559	6.387	-25.001	34.046	14.120	4.128	5.934	40.203

Note: To construct the 95 percent confidence set, we compute the minimum and maximum values of each parameter among the 95 percent of Metropolis-Hastings draws (after burning and thinning) with the highest values of the likelihood.

Table B.5: Metropolis-Hastings Draws of TU Cost Parameters

	μ				σ			
			95% confidence set				95% confidence set	
	Mean	SD	LB	UB	Mean	SD	LB	UB
TU 401	462.882	6.223	434.242	492.351	25.258	5.491	11.646	59.539
TU 402	409.170	8.409	364.553	445.384	34.828	7.342	19.305	75.865
TU 403	500.153	4.549	480.217	519.402	18.029	4.019	8.191	44.645
TU 404	619.820	58.410	438.540	812.952	274.968	47.589	166.877	439.077
TU 405	549.370	14.984	488.505	623.228	62.880	11.855	33.318	126.407
TU 501	470.523	13.027	421.373	523.690	53.055	10.943	29.483	114.743
TU 502	371.022	14.678	303.596	416.804	60.811	10.796	32.745	104.638
TU 503	404.424	11.086	357.965	447.487	46.099	8.700	25.910	90.902
TU 504	389.384	8.013	348.499	423.183	31.509	7.748	15.528	93.869
TU 505	400.846	9.514	367.954	442.166	38.464	7.857	17.821	97.961
TU 506	411.759	7.021	374.353	435.900	25.834	10.239	11.028	82.689
TU 507	414.557	8.962	374.139	452.098	36.079	7.895	19.096	87.309
TU 508	368.979	6.619	344.987	395.647	23.611	5.007	11.442	46.705
TU 510	385.638	18.056	311.150	455.317	82.265	15.674	43.593	155.060
TU 511	395.946	8.124	369.086	429.932	31.944	6.946	13.101	76.150
TU 901	466.919	11.248	416.666	512.511	49.072	12.149	15.326	123.650
TU 902	480.806	13.166	427.306	539.300	53.675	12.224	27.554	139.168
TU 903	521.769	23.808	439.945	630.150	98.422	30.706	23.153	185.951
TU 904	416.109	5.616	394.391	448.530	22.078	4.863	11.771	52.692
TU 905	433.060	12.728	382.972	482.855	54.647	10.097	31.840	98.462
TU 906	453.025	8.670	396.626	490.295	31.768	10.166	15.857	102.798
TU 907	465.834	10.842	429.383	517.135	43.403	10.430	13.868	95.346
TU 908	516.200	19.083	400.057	591.895	73.687	16.649	40.145	156.052
TU 909	503.912	19.666	416.887	584.630	88.452	15.876	50.673	168.220
TU 1201	547.574	17.934	477.209	648.365	69.930	14.976	37.236	160.578
TU 1331	370.662	14.950	312.350	429.718	64.083	11.224	32.686	125.590
TU 1332	348.331	16.872	259.191	418.628	72.816	14.634	41.662	186.140
TU 1333	329.089	16.026	247.694	434.639	70.122	12.591	39.913	140.245
TU 1334	386.891	28.422	260.902	483.792	126.246	20.673	70.418	225.348
TU 1335	369.221	11.345	327.185	409.970	47.260	8.444	24.269	89.210

Table B.5: Metropolis-Hastings Draws of TU Cost Parameters

	μ				σ			
			95% confidence set				95% confidence set	
	Mean	SD	LB	UB	Mean	SD	LB	UB
TU 1336	365.385	8.698	326.605	406.966	34.658	6.933	17.774	70.070
TU 1339	399.456	14.771	339.603	465.194	63.102	11.609	34.540	121.559
Incumbency shifter	-9.196	8.353	-43.462	17.798				

Note: To construct the 95 percent confidence set, we compute the minimum and maximum values of each parameter among the 95 percent of Metropolis-Hastings draws (after burning and thinning) with the highest values of the likelihood.

B.4.4 Type Distribution Estimates

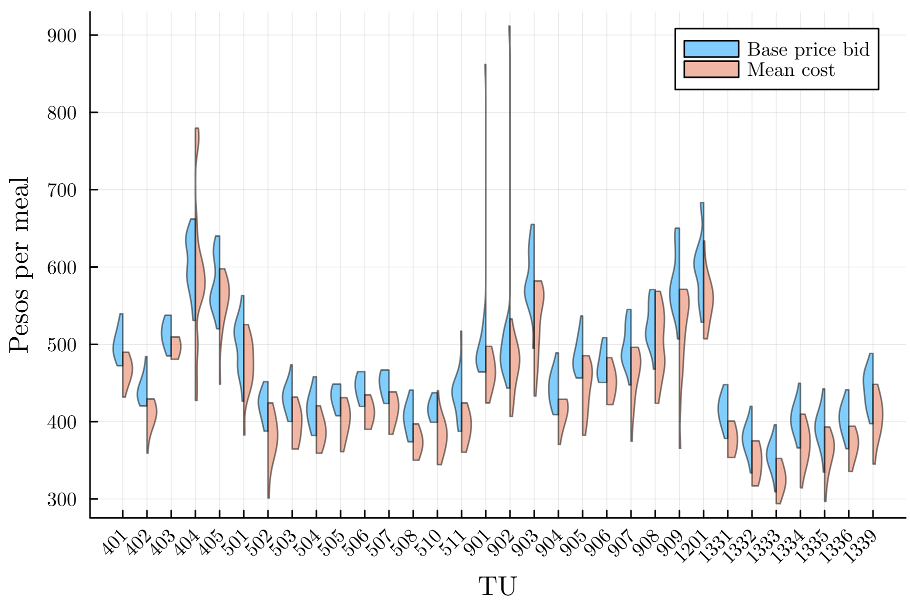
Across all bidders and all units, the mean unit cost is 434 pesos per meal, compared to the mean unit base price of 464 pesos per meal discussed in section 5.4.1. The average markup from unit cost to unit base price is 7.7 percent of the unit base price. Units on which bidders bid higher base prices are also higher-cost to supply. Figure B.2 compares the two distributions separately for each unit.

B.5 Welfare Analysis

B.5.1 Sampling from the Conditional Type Distribution

For our welfare analysis, we need draws from the estimated type distribution $F_{\Gamma_i|z_i}\left(\cdot \mid \hat{\theta}_\gamma\right)$ conditional on type bounds $\mathcal{G}_i$ for each bidder i. Obtaining these draws requires sampling from a multivariate normal distribution truncated to a polyhedron. While it is possible to sample from truncated univariate normals in closed form using inverse transform sampling, there is no equivalent procedure in higher dimensions. Furthermore, naive rejection sampling is computationally inefficient when the polyhedron's volume is small relative to the dimension of the space: most of the sampled draws from the unconditional distribution are discarded,

Figure B.2: Distributions of Firms' unit Base Prices and Costs



Note: These violin plots show the distributions of bidders' unit base prices $\tilde{\beta}_{ik}$ (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 $\mathcal{G}_i$.

so the number of unconditional draws required in order to obtain each conditional draw is high.

Instead, we sample bidder types γ_i via Gibbs sampling, drawing one dimension at a time as in appendix B.4.2. We run the Gibbs sampler for 100,000 iterations, discarding the first half as burn-in and thinning the remainder by a factor of five. We end up with $S_\gamma = 10{,}000$ draws of each bidder's cost parameters, which we use for the parametric results in Section 5.5.

B.5.2 Awarded Bid Cost and Markup Estimates

We present estimates of package costs and markups for each of the winning bids in table B.6.

Table B.6: Awarded Bid Cost and Markup Estimates

Firm	TUs	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	415.80	462.98	450.56	0.00	10.19	2.68
13	[905, 1334]	5.23	386.07	270.69	386.07	330.80	-0.00	29.89	14.32
16	[504, 505, 507]	6.42	391.58	-582.71	391.58	329.92	0.00	248.81	15.74
17	[403, 404, 502, 503, 904, 907, 1201, 1331]	20.91	405.64	304.40	404.98	379.93	0.16	24.96	6.34
19	[1332, 1333, 1336]	9.48	340.05	338.56	339.90	339.49	0.04	0.44	0.17
26	[501]	1.64	459.07	441.15	459.07	453.69	0.00	3.90	1.17
28	[506, 508, 510, 511]	10.22	381.66	373.63	381.66	379.62	0.00	2.10	0.53
36	[1335, 1339]	3.60	381.77	323.21	381.77	370.04	-0.00	15.34	3.07
47	[405]	2.08	536.59	444.01	460.39	447.98	14.20	17.25	16.51