Procurement Mechanisms for Differentiated Products

**Problem definition:** We consider the problem faced by a procurement agency that runs an auction-type mechanism to construct an assortment of differentiated products with posted prices, from which heterogeneous consumers buy their most preferred alternative.

**Academic / Practical Relevance:** Procurement mechanisms used by big organizations often take this form, including Framework agreements (FAs), that are widely used in the public sector.

**Methodology:** We use tools from mechanism design and auction theory to study the problem faced by the procurement agency and derive concrete practical recommendations.

**Results:** First, imposing different practical constraints such as linear pricing, we characterize the optimal buying mechanisms, which typically restrict the entry of close-substitute products to the assortment to increase price competition, without much damage to variety. Second, we use these optimal mechanisms as benchmarks to recommend improvements to the Chilean government procurement agency's implementation of FAs, used to acquire over US$2 billion worth of goods per year.

**Managerial Implications:** We devote the last section of our paper to explain in detail how our results and insights are changing the way FAs are implemented by the Chilean government.

1 Introduction

The Chilean government procurement agency (Dirección ChileCompra), our collaborator in this project, spends around US$2 billion yearly through framework agreements, approximately 20% of the total value of public procurement (Área de Estudios e Inteligencia de Negocios, Dirección ChileCompra 2017). A *framework agreement* (FA) is a procurement mechanism in which the central government procurement agency selects an assortment of differentiated products within certain category (e.g., computers) through competitive bidding in an auction mechanism. Then, public organizations, who have heterogeneous preferences over specific items within the category (e.g., over different computer models), buy their most preferred product from the selected assortment as needed. As an example, when buying laptops, a public school may want attractive graphics features while the Department of the Treasury may need high processing power. Similarly, organizations buying food may have different needs due to dietary constraints.

Even though each organization could in principle be in charge of their own purchases, the rationale is that, by managing the procurement process centrally, the procurement agency may be
able to exploit the purchasing power of a large buyer while providing variety to satisfy consumers’ heterogeneous needs. Therefore, it is not surprising that FAs have been widely adopted, not only in Chile, but worldwide: in 2010, the European Union awarded €80 billion using FAs in 2010, equal to 17% of the total value of public procurement in the EU (European Commision 2012). Further, this type of mechanisms are also extensively used in the private sector. For example, firms and universities typically use assortments of ‘selected’ suppliers and products from which their employees should buy from. Similarly, health plans maintain drug formularies —lists of prescription drugs available to enrollees for free or at a minimum co-pay— to help manage drug costs.

This paper is among the first in the literature to provide a formal analysis of this type of procurement mechanisms and improve our understanding of FAs. Our three main contributions are: (1) we introduce a model for the problem faced by the procurement agency to balance variety to account for consumers’ heterogeneous preferences and purchasing cost; (2) we characterize the optimal mechanisms under various realistic constraints; and (3) we use these results to study the design of simpler mechanisms that are commonly used in practical settings like the Chilean case. In addition to these contributions, the insights derived in this paper had a direct practical impact as they have led to concrete changes in the implementation of the Chilean government’s FAs. We describe these contributions in more detail next.

Our first main contribution is introducing a model that captures the following trade-off faced by a procurement agency when buying differentiated products. On one hand, consumers have heterogeneous preferences and, therefore, increasing product variety may increase consumer satisfaction. On the other hand, reducing the number of products bought may increase suppliers’ incentives to bid aggressively, so that their products have a better chance to be part of the small selection of items. We extend classical mechanism design and auction theory models to study this trade-off.

Using this model, we first analyze a centralized procurement mechanism, in which the buying agency decides the allocations of goods from suppliers to consumers. We characterize the optimal direct-revelation mechanism for a broad class of affine demand models, which are commonly used in competition models as they provide a reasonable balance between tractability and generality (see, e.g., Vives (2001)). The optimal mechanism illustrates how to introduce competition among close-substitute products to optimize the trade-off between variety and prices in terms of suppliers’ costs, products’ characteristics, and substitution patterns across products.

While the centralized mechanism just described is used in practice (e.g., the Chilean government uses a similar mechanism for large health-related purchases) and thus interesting on its own, we use it primarily as the basis to study a class of mechanisms incorporating two salient features of FAs.
First, FAs are mechanisms that select an assortment of suppliers offering different products at given unit prices, but cannot decide on how to allocate goods. The allocation is instead decentralized, decided by consumers based on the assortment and prices. Second, we impose the constraint that payments to suppliers must be implemented through linear prices with no upfront transfers, which is the payment structure broadly used in practice. Adding these realistic features introduces significant complexities in the characterization of the optimal mechanism and overcoming them is one of the main technical contributions of our work. In particular, the fact that the auctioneer cannot directly decide how to allocate demand across the products is a distinctive feature of our model relative to the traditional mechanism design problem introduced by Myerson (1981). Surprisingly, we show that, for a broad class of models considered in the paper, there is no performance loss associated to decentralizing the allocations and imposing linear pricing.

Our third main contribution is to leverage the characterization of the optimal decentralized mechanism to improve our understanding of the performance of first-price auction-type mechanisms, which are used in practice to award FAs. Using theoretical and numerical analysis, we show how auction rules that emulate the optimal mechanism by forcing close-substitute products to compete to be part of the menu can significantly decrease suppliers’ bids without much damage to variety.

Finally, in the last section we explain how the insights in this paper led to the redesign of ChileCompra’s food FA. An important observation that arises from studying their old FA design is that, because product definitions were narrow and auctions for different products were run independently, there was little to no competition at the auction stage. The redesign, which resulted from a collaboration between the authors, a research group in Universidad de Chile, and ChileCompra, focused on grouping products to generate more competition in the auction, while satisfying practical constraints. ChileCompra’s Director, Trinidad Inostroza, says: “The work in this paper was eye-opening for us. It made us realize the power of introducing competition at the auction stage to decrease prices while still providing adequate levels of variety. Having theory backing up this intuition together with numerical magnitudes convinced us to change our auctions for FAs.”

**Related literature.** Our work is closely related to some previous papers in procurement and regulation economics. Dana and Spier (1994) study how to allocate production rights to firms that have private cost information. An important insight of theirs is that the optimal market structure may depend on the firms’ bids, which is similar to our result that the optimal allocation depends on suppliers’ cost declarations. However, their auction only determines the market structure and lump-sum fees, whereas unit prices payed by consumers are determined by an exogenous competition
model. In contrast, our decentralized model captures two realistic features of FAs: (i) linear pricing; and (ii) the fact that these unit prices are endogenously determined by the mechanism.

Similarly, Anton and Gertler (2004) and McGuire and Riordan (1995) study the optimal mechanism with an endogenous market structure in a Hotelling model. However, unit prices are not part of the mechanism, and allocations are determined by the designer and not endogenously as in our decentralized case. Closer to our work, Wolinsky (1997) studies a spatial duopoly model where firms compete in both prices and quality. While it considers endogenous demands, the analysis restricts to solutions in which both firms have positive demands. In contrast, we are particularly interested in solutions in which some firms may be left out of the assortment to induce more competition.

Another stream of related work considering endogenous market structures is that of split-award auctions or dual sourcing (Chaturvedi et al. 2014, Li and Debo 2009, Elmaghraby 2000, Riordan and Sappington 1989, Anton and Yao 1989). However, these papers do not assume an underlying set of heterogeneous consumers as we do; instead, purchases are decided by the auctioneer.

Our work is also related to the operations literature studying assortment planning decisions (see e.g. Kök et al. (2009)). In these settings, decisions are made by one retailer that carries all products, and has full information regarding their unit costs. In our case instead, an assortment is built using an auction that elicits private cost information from many different suppliers.

Our analysis in Section 5 is closely related to Demsetz auctions (Demsetz 1968) which introduce competition for the market; Engel et al. (2002) also study a similar problem on a stylized model. This also relates to papers in group buying showing that committing to a single seller can be convenient for the group even if the members have heterogeneous preferences, as this can reduce buying prices (Dana 2012, Chen and Li 2013). However, these papers study complete information models; we extend their analysis to an auction setting with asymmetric cost information.

Finally, to the best of our knowledge, FAs are directly studied by only two prior papers. Albano and Sparro (2008) consider a complete-information Hotelling model with equidistant firms, where only the subset of suppliers with lowest bids are added to the assortment. Instead, we consider an incomplete information setting with a richer set of rules in which the assortment can depend on product characteristics. Gur et al. (2016) consider a model of FAs that studies the cost uncertainty faced by a supplier over the FA time horizon when selling a single-item, but does not consider multiple differentiated products nor heterogeneous consumers.

Overall, to the best of our knowledge, our work is the first to study optimal buying mechanisms in an asymmetric information setting, with an endogenous market structure, endogenous demand for differentiated products, and in which unit prices are determined by the mechanism.
2 Model

We introduce a model of procurement mechanisms for differentiated products. In our setting, the designer runs an auction-type mechanism to satisfy the demand arising from consumers for heterogeneous products provided by suppliers. Therefore, the actors in our model are (i) an auctioneer (or designer); (ii) suppliers (or agents); and (iii) consumers. We describe each of these next.

**Suppliers.** There is an exogenous set $N$ of $n$ potential suppliers indexed by $i$. Suppliers offer differentiated products that are imperfect substitutes to each other. The number of suppliers and the characteristics of their products are common-knowledge. We assume suppliers are risk-neutral, and seek to maximize expected profits. To simplify the exposition, we initially assume that each supplier offers exactly one product, so firms and products share the same indices. In a separate electronic companion we discuss the extension to the multi-product setting, and it is worth highlighting that our main results regarding optimal mechanism design hold under this extension.

Following the tradition in the auctions’ literature (see, e.g., Krishna (2009)), we assume that suppliers have production costs drawn independently from common-knowledge distributions, whose realizations are the private information of each supplier. Formally, supplier $i$ has a private cost $\theta_i \in \Theta_i$, associated to producing one unit of its product, where $\Theta_i$ is a finite set of strictly positive real numbers. We index the elements of $\Theta_i$, such that $\theta_i^j < \theta_i^k$ whenever $j < k$, for all $\theta_i^j, \theta_i^k \in \Theta_i$. We say that supplier $i$ is of type $\theta_i$ if his cost is $\theta_i$. Let $f_i$ be a probability mass function over $\Theta_i$, where $f_i(\theta_i)$ represents the probability that supplier $i$ is of type $\theta_i$. Let $F_i(\theta_i) = \sum_{k \leq j} f_i(\theta_i^k)$ be the cumulative probability distribution. Let $\Theta = \prod_i \Theta_i$ denote the type space. We use discrete distributions for technical convenience, as explained in Section 4. Because suppliers’ types are independent, the joint probability of $\theta = (\theta_1, \ldots, \theta_n)$ is equal to $f(\theta) = \prod_{i=1}^n f_i(\theta_i)$. We denote the probability that all suppliers other than $i$ have type $\theta_{-i}$ by $f_{-i}(\theta_{-i})$. We use boldfaces to denote vectors and matrices throughout the paper.

We assume that suppliers have constant marginal costs of production and do not face capacity constraints. Therefore, the products included in the assortment are always available and their production costs do not depend on the quantity demanded. These assumptions are typically reasonable in many settings we have in mind; for example, usually the quantities that suppliers sell through FAs represent only a small fraction of their total overall sales (Gur et al. 2016).

**Consumers.** We summarize aggregate consumer behavior and choices through a demand system. In the tradition of the assortment planning literature (e.g., Kök et al. (2009)) and the work
in oligopoly pricing (e.g., Tirole (1988)), we assume that aggregate demand functions are common knowledge and an input to our model. This assumption also seems reasonable in the contexts discussed in the introduction, as a demand system can typically be estimated using available historical data or consumer surveys (Ackerberg et al. (2006)).

In order to obtain analytical solutions to the mechanism design problem in Section 4, throughout the rest of the paper we restrict attention to a general class of affine demand models. These demand models capture a vast array of substitution patterns including both horizontal and vertical dimensions of differentiation, while being generally tractable. For these reasons, they have been extensively used in a variety of game-theoretic models within the operations literature (Cachon and Harker 2002, Allon and Federgruen 2007, Federgruen and Hu 2016).

Given a vector of prices $p \in \mathbb{R}^n$, let $d(p)$ denote the vector of demands, where the $i^{th}$ component represents the demand for supplier $i$. Traditionally, an affine demand function is one where the relation $d(p) = \alpha - \Gamma p$ holds for all $p \in \{p \in \mathbb{R} : \alpha - \Gamma p \geq 0\}$. Here, $\alpha \geq 0$ represents a quality component and $\Gamma$ is a matrix that captures substitution patterns across products. We assume that the products are substitutes, hence, $\Gamma_{ij} \leq 0$ for $i \neq j$, and that $\Gamma$ is symmetric and positive definite.

In addition, we assume that for any non-empty assortment and any vector of prices, total demand across all products in the assortment adds up to one. This assumption essentially amounts to imposing that there are no outside options and that the total demand is perfectly inelastic, which is a reasonable approximation for a variety of settings in public procurement where buying organizations cannot easily adjust the total quantity purchased based on prices, e.g. when buying medicines and school meals. However, as we briefly discuss in Section 4, our results extend to the case in which each product (or a subset of them) has an outside option.

Since we are particularly interested in solutions in which not necessarily all suppliers have positive demand, it is important to consider the extension of the affine demand specification introduced above to price vectors under which some products get zero demand, see Shubik and Levitan (1980) and Soon et al. (2009). We formalize this extension in our setting, where some suppliers can have zero demand and demands must add up to one, by assuming that a single ‘representative consumer’ maximizes consumer surplus (see Farahat and Perakis (2010)). The gross consumer surplus and consumer surplus for demand quantities $x$ and prices $p$ are, respectively, given by:

$$GCS(x) = c'x - \frac{1}{2}x'Dx,$$

and

$$CS(x, p) = GCS(x) - p'x,$$

where $D = \Gamma^{-1}$ and $c = \Gamma^{-1}\alpha$ have been renamed to avoid burdensome notation. Consumer sur-
plus is a quasi-linear function of prices, which will be important to solve for the optimal mechanism.

The demand function is defined as the solution of the representative consumer’s maximization problem. (An alternative way of micro-founding an aggregate demand system is to start from a discrete choice model that describes individual consumption decisions; see Armstrong and Vickers (2014) for a discussion on the affine demand models used in this paper.) Hence, for any \( p \in \mathbb{R}^n \), let \( d(p) \) be defined as the solution of:

\[
\max_x \ CS(x, p) = c'x - \frac{1}{2} x'Dx - p'x \quad \text{s.t.} \quad 1'x = 1, \quad x \geq 0.
\]

Clearly, Problem (2) has a unique solution for every \( p \in \mathbb{R}^n \); hence, the demand function \( d(p) \) is well defined. Also, depending on the price vector, the solution of this maximization problem may set some of the demand quantities equal to zero, hence leaving some suppliers without demand.

Further, for a set of suppliers \( Q \subseteq N \) we can extend Problem (2) by imposing the additional constraints \( x_i = 0, \forall i \notin Q \), i.e. a-priori selecting the set of potential firms to buy from. We denote by \( d(Q, p_Q) \) as the solution of this extended problem where \( p_Q \) denotes the elements of the price vector corresponding to the indices in \( Q \). Under this notation, observe that \( d(N, p_N) = d(p) \).

For a given price vector \( p \) and \( Q \subseteq N \), we let \( Q' \subseteq Q \) be the set of active firms, i.e. those with \( d_i(Q, p_Q) > 0 \). In the separate electronic companion, we show that the positive part of \( d(Q, p_Q) \) is an affine function of the prices of the set of active suppliers only and, therefore, can be written as \( d(Q, p_Q) = \alpha - Gp_{Q'} \). (Even in the case \( Q = Q' = N \), \( \alpha \) and \( G \) are not equal to \( \alpha \) and \( G \), respectively, because of the additional constraint \( 1'x = 1 \).) We also show that demand for a product is weakly decreasing in its own price and increasing in others’ prices. Importantly, we note that cross price elasticities change as a function of the set of active firms.

**Auctioneer.** The role of the auctioneer is to design a mechanism to satisfy the purchasing needs of the heterogeneous consumers. The auctioneer is risk-neutral and her objective is to maximize expected gross consumer surplus (as defined in Eq. (1)) minus the expected payments to suppliers. Note that if transfers to suppliers are implemented using linear pricing without upfront fees (the transfer to each supplier equals his price times the quantity demanded), then the objective is to maximize consumer surplus; this case is relevant in practice and discussed in detail in Section 4.

This is an adequate objective in the applications described in the introduction, as it incorporates both variety and cost considerations. In particular, the aggregate value derived by consumers from variety is captured by the gross consumer surplus term, while the total procurement cost is captured by the transfers to suppliers. Having said that, our approach could also be extended to maximize
social welfare instead, albeit with some non-trivial modifications.

As we discussed in the introduction, we consider two classes of procurement mechanisms that we call centralized and decentralized. In both settings, the rules of the mechanism are common-knowledge. In a centralized procurement setting, the auctioneer runs a mechanism to decide the fraction of the aggregate demand that will be satisfied by each supplier, and their appropriate compensations. In this case, the auctioneer decides how to distribute the goods among the different consumers, that is, the auctioneer acts like a central planner who can decide how to allocate goods. While the auctioneer has the ability to allocate demand, he does not have access to suppliers’ private information and thus the mechanism is used for price discovery. The optimal centralized mechanism is formulated and analyzed in Section 3.

In contrast, in a decentralized setting, the auction is run to construct the menu of products based on suppliers’ bids. The menu consists of a subset of suppliers and unit prices for their products. Once the menu is fixed, consumers are free to choose which product in the menu to purchase. That is, the main difference with the centralized setting is that the auctioneer cannot directly determine the resulting allocations; allocations are determined by consumers through the underlying demand system. In other words, the quantities purchased by consumers must be consistent with the demands defined by the solution to Problem 2. As it will become clear once both problems are formally stated, the centralized problem can be seen as a relaxation of the decentralized problem: in the centralized setting the designer determines the allocations and constraints involving the demand system are ignored. We describe the decentralized mechanism in Section 4.

To conclude, we note that centralized mechanisms are used in some situations, and thus interesting on their own. For instance, Cenabast, a procurement organization within the Chilean government which has its own budget and makes large health-related purchases to distribute to different organizations, runs a centralized auction taking into account variety considerations. However, in the context of this paper, our focus is in understanding the decentralized setting as it better corresponds to the way FAs are implemented. Thus, the main purpose of the centralized mechanism in this paper is to provide a benchmark on what is achievable (recall that it can be seen as a relaxation of the decentralized one), and useful insights on the optimal market structure and thus on the trade-off between variety and payments to suppliers.
3 Centralized Procurement

3.1 Mechanism Design Problem Formulation

We provide a mechanism design formulation for the auctioneer’s problem, considering Bayes Nash implementation. By invoking the revelation principle, we restrict attention to direct revelation mechanisms without loss of optimality. Hence, for given cost declarations, the designer selects an allocation of the total demand across the different suppliers, as well as their appropriate compensations. Formally, a direct revelation mechanism can be specified by (a) the allocation functions $x_i : \Theta \rightarrow [0, 1]$, where $x_i(\theta)$ is the quantity allocated to supplier $i$ when cost declarations are $\theta$, and (b) the transfer functions $t_i : \Theta \rightarrow [0, 1]$, where $t_i(\theta)$ denotes the payment to supplier $i$ when cost declarations are $\theta$. Let $x = (x_1, \ldots, x_n)$ and $t = (t_1, \ldots, t_n)$.

In the optimal mechanism design problem, the designer maximizes its objective (in our case, expected gross consumer surplus minus expected transfers) subject to the usual constraints in mechanism design theory: incentive compatibility (IC), individual rationality (IR), and feasibility of allocations (Feas). To write these constraints, we define the interim expected utility for supplier $i$ of type $\theta_i$ and report $\theta_i'$ as follows:

$$U_i(\theta_i' | \theta_i) = \sum_{\theta_{-i} \in \Theta_{-i}} f_{-i}(\theta_{-i}) (t_i(\theta_i'| \theta_{-i}) - \theta_i x_i(\theta_i'| \theta_{-i})), \quad (3)$$

where $\theta_{-i}$ is the report of supplier $i$’s competitors. Now, the auctioneer’s optimal mechanism design problem can be formulated as follows:

$$\text{[Cent]} \quad \max_{x,t} \quad \mathbb{E}_\theta \left[ \text{GCS}(x(\theta)) - \sum_{i=1}^n t_i(\theta) \right]$$

s.t.  
$$U_i(\theta_i | \theta_i) \geq U_i(\theta_i' | \theta_i) \quad \forall i \in N, \forall \theta_i, \theta_i' \in \Theta_i \quad \text{ (IC)}$$

$$U_i(\theta_i | \theta_i) \geq 0 \quad \forall i \in N, \forall \theta_i \in \Theta_i \quad \text{ (IR)}$$

$$\sum_{i \in N} x_i(\theta) = 1 \quad \forall \theta \in \Theta, \quad x_i(\theta) \geq 0 \quad \forall i \in N, \theta \in \Theta \quad \text{ (Feas)}$$

This problem only differs from the classic mechanism design formulation in the objective function: while in the latter expected transfers are minimized, in the former we maximize expected gross consumer surplus minus transfers. Therefore, the optimal solution to Cent can be obtained by extending standard arguments based on the envelope theorem (Myerson 1981) adapted for the setting of discrete distributions (Vohra 2011) to determine which IC constraints bind.
Analogously to the setting of continuous cost distributions, we introduce the following definition of the virtual cost function for cost distributions with discrete support (see Vohra (2011)).

**Definition 3.1.** For $\theta_i \in \Theta_i$, let $\rho_i(\theta_i) = \max\{\theta' \in \Theta_i : \theta' < \theta_i\}$, that is, $\rho_i(\theta_i)$ is the predecessor of $\theta_i$ in $\Theta_i$. (If $\theta_i$ is the lowest cost in the support, we define $\rho_i(\theta_i) = \theta_i$.) Let $v_i(\theta_i) = \theta_i + \frac{F_i(\rho_i(\theta_i))}{f_i(\theta_i)}(\theta_i - \rho_i(\theta_i))$ be the virtual cost function of supplier $i$. Let $v(\theta) = (v_1(\theta_1), \ldots, v_n(\theta_n))$.

We make the standard regularity assumption in mechanism design that we keep throughout:

**Assumption 3.1.** The function $v_i(\theta_i)$ is strictly increasing for all $i \in N$.

Finally, we also define the *interim expected allocations* and *interim expected transfers* as follows:

$$X_i(\theta_i) \equiv \sum_{\theta_{-i} \in \Theta_{-i}} f_{-i}(\theta_{-i})x_i(\theta_i, \theta_{-i}), \quad T_i(\theta_i) \equiv \sum_{\theta_{-i} \in \Theta_{-i}} f_{-i}(\theta_{-i})t_i(\theta_i, \theta_{-i}).$$

The optimal solution to Problem $Cent$ can be characterized as follows:

**Proposition 3.1.** Suppose that $(x, t)$ satisfy the following conditions:

1. The allocation function satisfies for all $\theta \in \Theta$,

$$x(\theta) \in \arg\max \, CS(x(\theta), v(\theta)) \quad \text{s.t.} \quad \sum_{i=1}^n x_i(\theta) = 1, \quad x_i(\theta) \geq 0 \quad \forall i \in N. \quad (4)$$

2. Interim expected allocations are monotonically decreasing for all $i \in N$, that is, $X_i(\theta) \geq X_i(\theta')$ for all $\theta, \theta' \in \Theta_i$ such that $\theta \leq \theta'$.

3. Interim expected transfers satisfy for all $i \in N$ and $\theta_i^j \in \Theta_i$:

$$T_i(\theta_i^j) = \theta_i^j X_i(\theta_i^j) + \sum_{k=j+1}^{\lfloor |\Theta_i| \rfloor} (\theta_i^k - \theta_i^{k-1})X_i(\theta_i^k) \quad (5)$$

Then, $(x, t)$ is an optimal mechanism for the centralized procurement problem $Cent$.

The proof can be found in Appendix A. Although the method of analysis of the centralized mechanism is quite standard, we have not seen such a transparent characterization of the trade-off between variety and payment to suppliers in the literature before. Condition (1) in Proposition 3.1 states that, for each $\theta \in \Theta$, the optimal vector of allocations $x(\theta)$ coincides with the demand functions defined by Problem (2) when unit prices are equal to virtual costs. This follows from classic
mechanism design arguments, i.e., the equilibrium ex-ante expected payment that the auctioneer makes to a bidder is equal to the ex-ante expectation of the virtual cost times the allocation.

Finally, note that while the optimal demands are fully characterized, the optimal transfers are not. The only constraint imposed on transfers at optimality is over expected transfers (condition (3) in Proposition 3.1). This freedom in the definition of optimal prices becomes useful later on, when we characterize the optimal solution to the decentralized problem under linear pricing.

### 3.2 Optimal Centralized Mechanism for Hotelling Demand Model

To fix ideas and provide intuition, we first discuss in detail the structure of the optimal centralized mechanism when the consumer demand is given by a Hotelling model; we then discuss the general affine demand model in Section 3.3.

We start by briefly discussing a general Hotelling demand model with an arbitrary number $n$ of suppliers located at $0 < \ell_1 < \ell_2 < \ldots < \ell_n \leq 1$; the location represents the horizontal characteristic of the product offered by the supplier relative to the product space. The closer two suppliers are in the product space, the closer substitutes the products offered by them are. A continuum of consumers, all of whom must buy one unit of product, are distributed on the product space. To simplify the exposition, we assume that consumers are uniformly distributed. The utility consumer $j$ obtains from buying the product offered by $i$ is given by: $u_{ji}(p_i) = - (\delta |\ell_i - \ell_j| + p_i)$, where $\delta$ is the transportation cost and $\ell_j$ is the position of consumer $j$ in the unit line. Using these utility functions we can characterize the demand function and the centralized optimal solution.

Suppose that suppliers have fixed unit prices $p = \{p_i\}_{i \in \mathbb{N}}$. It is easy to see that supplier $i$ will have positive demand if and only if $i$ is the preferred choice for the consumer located at $\ell_i$. Therefore, the set of active suppliers is given by $Q(p) = \{i \in \mathbb{N} : p_i \leq \min_{k \neq i} \{p_k + \delta |\ell_k - \ell_i|\}\}$, where we abused notation to make the set depend on prices instead of costs. In addition, two consecutive active suppliers $i$ and $j$ split the segment between $\ell_i$ and $\ell_j$ proportionally to their prices: $i$ obtains $\frac{p_i - p_j + \delta |\ell_i - \ell_j|}{2\delta}$ and $j$ the rest. (The demand equations can be easily derived by determining the location of an indifferent consumer between two active neighboring suppliers.)

The optimal solution to problem $\text{Cent}$ under the Hotelling model is intuitive. By Proposition 3.1, for any cost realization $\theta$, the optimal allocations are given by the Hotelling demands when prices are equal to the vector of virtual costs $v(\theta)$. Therefore, for a given $\theta \in \Theta$, the optimal assortment is given by: $Q(\theta) = \{i \in \mathbb{N} : v_i(\theta_i) - v_j(\theta_j) \leq \delta |\ell_j - \ell_i| \quad \forall j \in \mathbb{N}\}$, which corresponds to the definition of the set of active suppliers when prices are replaced by virtual costs.

To illustrate the result, consider a simple Hotelling model with two suppliers sharing the same
cost distribution, and positioned at the extremes of the unit interval (so \( \ell_1 = 0, \ell_2 = 1 \)). Let \( \theta_1 \) and \( \theta_2 \) be the cost realizations of suppliers 1 and 2, respectively. In this case, the centralized problem yields an optimal allocation characterized by: (1) if \( \delta > |v(\theta_1) - v(\theta_2)| \), the demand is split between the two suppliers with \( x_1 = (v(\theta_2) - v(\theta_1) + \delta)/(2\delta) \) and \( x_2 = (v(\theta_1) - v(\theta_2) + \delta)/(2\delta) \); and (2) if \( \delta < |v(\theta_2) - v(\theta_1)| \), all the demand is awarded to the supplier with the lowest cost realization.

As illustrated, an important feature of the centralized optimal solution is that the decision of whether to split or not the demand depends on the cost realizations. If the transportation cost is small relative to the differences in virtual costs, then it is optimal to have a unique supplier, the one with the lowest virtual cost. In this case, it is worth paying the cost of having less variety in the assortment with the upside of decreasing the expected payments to bidders. In contrast, if the transportation cost is high relative to the differences in virtual costs, the demand is split between both suppliers to increase variety. In this sense, the optimal solution to the centralized problem optimizes the trade-off between variety and payments to suppliers: by restricting the entry to the assortment in some scenarios, the auctioneer can reduce expected payments while still providing incentives for truthful cost revelation. This insight generalizes to the case with more than two suppliers. In particular, if two products are close substitutes (i.e., \( \delta|\ell_j - \ell_i| \) is relatively small) then the auctioner will not purchase the product with the highest virtual cost. On the other hand, when two products are not close substitutes (i.e., \( \delta|\ell_j - \ell_i| \) is relatively big), then the (virtual) cost of one product is less likely to affect whether the other product is included or not in the assortment.

### 3.3 Optimal Centralized Mechanism for General Affine Demand Models

We now turn our attention to the more general affine demand models introduced in Section 2, that allow us to combine both vertical and horizontal sources of differentiation. Recall that for a general affine demand model, the demand functions are obtained by solving Problem (2). To gain some intuition, we start discussing a simple example with two suppliers.

**Example 3.1.** We consider a duopoly where \( \alpha = (a_1, a_2) \) and \( \Gamma = \left( \frac{r_1}{-\gamma}, \frac{r_2}{\gamma} \right) \), with all the parameters positive and with \( r_1 + r_2 \geq 2\gamma \). Under these parameters, we have \( D = \frac{1}{r_1 r_2 - \gamma^2} \left( \frac{r_2 a_1 + \gamma a_2}{r_1 a_2 + \gamma a_1} \right) \) and \( c = \frac{1}{r_1 r_2 - \gamma^2} \left( \frac{r_2 a_1 + \gamma a_2}{r_1 a_2 + \gamma a_1} \right) \). For any given \( p \), the demand functions \( d(p) \) are given by:

\[
d_i(p) = \max \left\{ 0, \min \left\{ \frac{(r_j - \gamma) a_i - (r_i - \gamma) a_j + r_i - \gamma - (r_i r_j - \gamma^2)(p_i - p_j)}{r_i + r_j - 2\gamma}, 1 \right\} \right\}, \quad i, j \in \{1, 2\}
\]

Recall that, for a given cost realization \( \theta \), the optimal allocations in the centralized problem equal the demand characterized above with prices equal to the vector of virtual costs \( v(\theta) \). To
illustrate, we discuss the structure of the optimal solution in Example 3.1, focusing on supplier 1. For a given $\theta$, he will be in the assortment $(d_1(\theta) > 0)$ if and only if $(r_1 r_2 - \gamma^2) (v_1(\theta) - v_2(\theta)) \leq (r_2 - \gamma) a_1 - (r_1 - \gamma) a_2 + r_1 - \gamma$. Therefore, for him to be active, the difference in virtual costs must be bounded by a quantity that is increasing in the ‘normalized’ quality difference $(r_2 - \gamma) a_1 - (r_1 - \gamma) a_2$. Hence, the larger this difference (e.g., if $a_1$ grows), the larger the difference in virtual costs allowed to keep supplier 1 active. Note that, similarly to the Hotelling model, the centralized optimal solution may restrict the entry of a supplier to the assortment to decrease expected payments. The structure of the centralized optimal allocation generalizes to the case of more products.

4 Decentralized Procurement

The objective of this section is to characterize the optimal mechanism in the auctioneer’s decentralized problem, that more closely resembles FAs. In the decentralized setting, for given cost declarations, the auctioneer selects a menu which consists of an assortment of products (or suppliers) and their unit prices. Purchasing decisions are decentralized: based on the products and prices in the menu, consumers decide which products to buy through the demand system (see Section 2).

In this setting, the payments to the suppliers need not be linear in prices (i.e., need not to be equal to suppliers equal unit price times quantity sold) as, in principle, the auctioneer might choose to implement a more general payment structure. However, general payment structures rarely arise in the real-world settings we are trying to capture. Therefore, throughout the rest of the paper, we purposely assume linear pricing and no upfront fees in our decentralized model.

Imposing linear pricing introduces significant technical challenges when characterizing the optimal mechanism, and a critical portion of this section is devoted to overcoming these challenges. In contrast, we argue that characterizing the optimal mechanism under a more general payment structure such as a two-part tariff is significantly easier (see Remark 4.1). However, linear pricing is an important practical constraint that any realistic model of FAs should incorporate. Gian Luigi Albano, a world-known expert in FAs, mentioned in a private communication that “In 99% of cases, I have just come across linear pricing” as the method to implement transfers to suppliers; see Chapter 6 in Albano and Nicholas (2016) for a summary of FAs’ implementations in different countries. (Further, regarding the procurement mechanism, he mentions that “all I have seen in eleven years is first-price auctions,” which will be the subject of Section 5.) Notably, we show that for a broad class of models, there is not any performance loss associated to using linear pricing.
4.1 Linear Pricing Mechanism Design Problem Formulation

We consider Bayes Nash implementation and restrict attention to direct revelation mechanisms without loss of optimality. A decentralized direct revelation mechanism can be specified by (a) the ‘assortment’ functions \( q_i : \Theta \to \{0, 1\} \) that are equal to 1 if and only if supplier \( i \) is included in the assortment when cost declarations are \( \theta \); and (b) the price functions \( p_i : \Theta \to \mathbb{R} \), where \( p_i(\theta) \) is the unit price for the item offered by supplier \( i \) when cost declarations are \( \theta \). Note that this formulation allows for multiple suppliers to be in the menu. We define \( q = (q_1, \ldots, q_n) \) and \( p = (p_1, \ldots, p_n) \). For given cost declarations \( \theta \), the menu is given by \( (q(\theta), p(\theta)) \). Analogously to the centralized setting, let \( x_i : \Theta \to [0, 1] \) denote the allocation functions, i.e., \( x_i(\theta) \) is the quantity allocated to supplier \( i \) when cost declarations are \( \theta \). Let \( x = (x_1, \ldots, x_n) \). In the decentralized optimal mechanism design problem, the auctioneer maximizes expected consumer surplus (as given by Eq. (1)) subject to incentive compatibility (IC), individual rationality (IR), feasibility of allocations (Feas), plus an extra set of constraints that relates allocations to the demand system (explained in detail below).

As in the centralized setting, the IC and IR constraints can be expressed in terms on the interim expected utilities. Using the linear-pricing assumption, the interim expected utility for supplier \( i \) of type \( \theta_i \) and report \( \theta'_i \) defined in Eq. (3) can be reexpressed

\[
U_i(\theta'_i|\theta_i) = \sum_{\theta_{-i} \in \Theta_{-i}} f_{-i}(\theta_{-i}) \left( p_i(\theta'_i, \theta_{-i}) - \theta_i \right) x_i(\theta'_i, \theta_{-i}).
\] (6)

In addition, we must include constraints to ensure that the allocations are consistent with the underlying demand system (Demand). That is, for every vector of cost realizations \( \theta \), the consumer demand associated to the menu \( (q(\theta), p(\theta)) \) is determined by the underlying demand system obtained by solving (2) (with the additional constraints that \( x_i = 0 \) if \( q_i = 0 \)). This is in sharp contrast with the centralized mechanism and classical mechanism design theory, in which the designer selects both a payment function and an allocation function. In our problem, the designer does not directly select the allocation function; instead, he chooses an assortment and unit prices and, given these, allocations are endogenously determined through the demand system. As it will become apparent in Section 4.2, the demand constraints imposed on the allocations together with our linear pricing structure introduce significant additional complexities when solving for the optimal mechanism.
The auctioneer’s optimal mechanism design problem can now be formulated as follows:

\[
[DecLin] \quad \max_{q,p,x} \mathbb{E}_\theta[CS(x(\theta), p(\theta))] \\
\text{s.t.} \quad U_i(\theta_i | \theta_i) \geq U_i(\theta'_i | \theta_i) \quad \forall i \in N, \forall \theta_i, \theta'_i \in \Theta_i \quad \text{(IC)} \\
U_i(\theta_i | \theta_i) \geq 0 \quad \forall i \in N, \forall \theta_i \in \Theta_i \quad \text{(IR)} \\
\sum_{i \in N} x_i(\theta) = 1 \quad \forall \theta \in \Theta, \quad x_i(\theta) \geq 0 \quad \forall i \in N, \forall \theta \in \Theta \quad \text{(Feas)} \\
x_i(\theta) = d_i(q(\theta), p(\theta)) \quad \forall i \in N, \forall \theta \in \Theta, \quad \text{(Demand)}
\]

where for each \( \theta, d(q(\theta), p(\theta)) \) corresponds to the solution of (2) with the additional constraints that \( x_i = 0 \) if \( q_i = 0 \). Note that we abused notation to also denote by \( q(\theta) \) the set of suppliers that are in the assortment given costs \( \theta \). In the next section we discuss the solution to the optimal mechanism design problem \( DecLin \) under affine demand models. As demands are obtained by maximizing consumer surplus and the auctioneer seeks to maximize expected consumer surplus, it might appear that the demand constraints are redundant. In Section 4.2 it will become clear that this is not the case because of the additional IC constraints.

### 4.2 Solution Approach for the Linear Pricing Decentralized Mechanism

The optimal mechanism design problem \( DecLin \) is a mixed integer program that takes a demand model as an input. Even after relaxing the integrality of the variables \( q \), the program is typically non-convex because demand equations are often non-linear, even in simple cases such as the Hotelling model specified below. (Although demand for active firms is linear in prices, the overall demand function also specifies the set of active firms for a given vector of prices, and this function is typically non-linear.) Moreover, the presence of demand constraints prevents us from directly applying the standard mechanism design arguments used in the centralized case; under these additional constraints it is not possible to tell a-priori which IC constraints bind in the optimal solution.

Our solution approach relies on relaxing these demand constraints and solving the centralized problem, which admits an analytical solution as discussed in Section 3. We then use this solution to construct an optimal solution to the decentralized problem. In other words, we show the existence of prices that allow us to decentralize the solution to the centralized problem.

Formally, note that if we relax the demand constraints in \( DecLin \) and rewrite the interim utilities in terms of total transfers \( t \) by setting \( t_i(\theta) = p_i(\theta)x_i(\theta) \), then this problem is equivalent to the centralized problem. Because problem \( Cent \) is a relaxation of \( DecLin \), the optimal objective of
the former is an upper bound on the optimal objective of the latter. The next corollary provides necessary and sufficient conditions under which DecLin attains the optimal objective of Cent.

**Corollary 4.1.** Let \((x, T)\) be the unique optimal solution to the centralized problem Cent, where \(T\) denotes the vector of interim expected transfers. Define

\[
q_i(\theta) = 1 \text{ if and only if } x_i(\theta) > 0, \forall i \in N, \theta \in \Theta. \tag{7}
\]

Suppose that for all \(\theta \in \Theta\), there exist prices \(p(\theta)\) such that

\[
x_i(\theta) = d_i(q(\theta), p(\theta)) \quad \forall i \in N, \forall \theta \in \Theta, \text{ and}
\]

\[
T_i(\theta_i) = \sum_{\theta_{-i} \in \Theta_{-i}} p_i(\theta_i, \theta_{-i})x_i(\theta_i, \theta_{-i})f_{-i}(\theta_{-i}), \quad \forall i \in N, \forall \theta_i \in \Theta_i. \tag{9}
\]

Then, the optimal objective of DecLin is equal to the optimal objective of Cent. Moreover, an optimal solution of DecLin is given by \((q, p)\) characterized by Eqs. (7), (8), and (9), and the corresponding optimal allocation \(x\) of Cent. Furthermore, the optimal objective of DecLin is equal to the optimal objective of Cent if and only if such solution \((q, p)\) exists.

The corollary suggests the following approach to solving the optimal decentralized mechanism design problem. First, solve the centralized problem, which can be viewed as a relaxation of the decentralized one. Then, find unit prices that allow to decentralize the optimal solution by: (1) making the aggregate demands under such prices agree with the centralized optimal allocations; and (2) satisfying the individual rationality and incentive compatibility constraints for suppliers' truthful revelation of information through the interim expected transfers. This is at the heart of the technical challenge in solving for the optimal decentralized mechanism: under a linear pricing structure we only have one instrument (unit prices) to accomplish these two tasks, and showing the existence of such prices is non-trivial. The above discussion also highlights why the demand constraints are not redundant in the presence of the IC constraints.

**System of Linear Equations.** Corollary 4.1 introduces a system of linear equations (given by Eqs. (8) and (9)) that unit prices must satisfy for a solution to the decentralized problem to achieve the centralized optimal objective. Equations (8) require that prices induce the optimal allocations \(x\) of Cent through the demand system. We denote by \(Q(\theta)\) the set of active suppliers, those with strictly positive demands in the centralized optimal solution. Hence, using Proposition 3.1, we need to find unit prices such that \(d_i(N, v(\theta)) = d_i(q(\theta), p(\theta)), \) for all \(i \in Q(\theta)\) and \(\theta \in \Theta.\)
As the demand function is assumed to be affine in prices, these equations yield linear constraints in prices. Note that the equations are linear because they require to find prices to generate a given vector of demands for active firms. Equations (9) require that unit prices $p$ induce the interim expected transfers $T_i$ in the optimal solution of $Cent$ —that is, the solution is individually rational and incentive compatible. Given an optimal solution to $Cent, (x, T)$, these equations are also linear in prices. In particular, once the constraints in Eqs. (8) are imposed, the allocations are fixed and equal to the optimal allocations of $Cent$; therefore, Eqs. (9) are linear in unit prices.

By the observations above, verifying whether $OPT(DecLin) = OPT(Cent)$ (here, $OPT(P)$ denotes the optimal value of problem $P$) for affine demand models is equivalent to establishing whether the linear system of equations defined by Eqs. (8) and Eqs. (9) admits a solution. (Assuming discrete types allows us to work with finite dimensional system of equations and to use finite-dimensional linear algebra. In the continuous type setting, we would have to deal with an infinite dimension space for price variables, and the results would be more technically involved.) In Appendix B we describe the system of equations in detail. In the remainder of this section, we show that (under additional mild conditions) we can guarantee that this system is consistent and admits a solution and, therefore, we can characterize the optimal mechanism.

**More general payment structures.** As mentioned in the introduction to this section, one could also consider more general payment structures to compensate suppliers. Perhaps a sensible mechanism would be to use a two-part tariff, in which the auctioneer receives (or pays) a fix transfer from every firm participating in the assortment and, in addition, each firm receives a linear transfer from consumers (equal to the unit prices set by the mechanism times the allocations). In that case, the problem can be easily solved by exploiting the solution to the centralized mechanism as follows.

**Remark 4.1.** In contrast to the setting under linear pricing, the solution to the decentralized problem could be vastly simplified under a two-part tariff. Intuitively, we can use the unit-prices to decentralize the optimal consumption and then the upfront fee to satisfy the IC constraints.

Formally, define the upfront payment functions $y_i : \Theta \to \mathbb{R}$, where $y_i(\theta)$ is the upfront payment received (or given) by supplier $i$ when cost declarations are $\theta$. The objective of the auctioneer can be rewritten as: $\mathbb{E}_\theta \left[ CS(x(\theta), p(\theta)) - \sum_{i \in N} y_i(\theta) \right] = \mathbb{E}_\theta \left[ GCS(x(\theta)) - \sum_{i \in N} (x_i(\theta)p_i(\theta) + y_i(\theta)) \right]$, and suppliers’ interim utilities as: $U_i(\theta_i|\theta) = \sum_{\theta_{-i} \in \Theta_{-i}} f_{-i}(\theta_{-i}) \left( (p_i(\theta_i, \theta_{-i}) - \theta_i) x_i(\theta_i, \theta_{-i}) + y_i(\theta_i, \theta_{-i}) \right)$. Note that, by defining $t_i(\theta) = x_i(\theta)p_i(\theta) + y_i(\theta)$, both the definitions of the objective and the interim utilities agree with those of the centralized problem defined in Section 3. Therefore, to solve for the optimal two-part tariff mechanism we can first solve the centralized mechanism to
obtain the optimal allocations and transfers \((x^*(\theta), t^*(\theta))\). Then, find prices \(p(\theta)\) that implement allocations \(x^*(\theta)\) by solving the associated linear system. Finally, define \(y_i(\theta) = t^*_i(\theta) - x^*_i(\theta)p_i(\theta)\).

The above remark establishes that, in fact, the centralized mechanism can in principle be implemented using a decentralized mechanism with a more general payment structure. However, we remind the reader that linear pricing is a prevalent practice in the environments we are trying to capture, and thus an important operational constraint that we think must be included. Further, we will show that for a broad class of models there is no optimality loss associated to linear pricing.

### 4.3 Optimal Linear Pricing Decentralized Mechanism for Hotelling Demand

We now provide conditions under which we can construct the optimal solution to the linear pricing decentralized mechanism by following the approach described in the previous section. Recall that this solution will have the same intuitive interpretation as the centralized solution, as the assortment, allocations, and expected payments agree.

We start with the Hotelling model. To that end, consider the optimal solution to problem \(\text{Cent}\), where for a cost realization \(\theta\), the allocations are given by the Hotelling demands when prices are equal to the vector of virtual costs. Let \(q\) be defined as in Corollary 4.1, that is, \(q_i(\theta) = 1\) if \(i \in Q(\theta)\), and \(q_i(\theta) = 0\) otherwise, where \(Q(\theta)\) is the set of active suppliers in the centralized optimal solution under profile \(\theta\). Recalling the form of the Hotelling demands and the optimal allocations of \(\text{Cent}\), it should be clear that the constraints given by Eqs. (8) can be summarized as:

\[
p_{\theta(i)}(\theta) - p_i(\theta) = v_{\theta(i)}(\theta_{\theta(i)}) - v_i(\theta_i) \quad \forall \theta \in \Theta, i \in Q(\theta), i \text{ is not the rightmost supplier,} \tag{10}
\]

and \(\theta_{\theta(i)}\) denotes the successor of \(i\) in the set \(Q(\theta)\) (i.e., \(\theta_{\theta(i)} = \min\{j \in Q(\theta) : j > i\}\)). In words, the difference in prices between adjacent active suppliers must be equal to the difference in their virtual costs. These constraints will implement the optimal allocations of \(\text{Cent}\) using prices \(p(\theta)\). These \(n - 1\) equations impose constraints over the price differences but not actual prices, a feature that we will exploit to make sure we also satisfy the constraints given by Eq. (9), that is, that expected transfers agree with the optimal ones.

Unfortunately, even when demand is given by the Hotelling model, the optima of both problems might not always agree. (Due to lack of space, an example is deferred to the electronic companion.) Therefore, we next focus on providing sufficient conditions under which \(\text{OPT}(\text{Cent}) = \text{OPT}(\text{DecLin})\). The following theorem establishes that, under additional mild conditions, both optima will agree and thus the optimal decentralized mechanism can be characterized.
Theorem 4.1. Consider the general Hotelling model in which suppliers have arbitrary locations and costs distributions. Let $c^* = \min_{1 \leq i \leq n-1} (\ell_{i+1} - \ell_i)$. Suppose that the following conditions are simultaneously satisfied:

1. There is at least one profile $\theta \in \Theta$ such that $|v_i^{i+1}(\theta_{i+1}) - v_i(\theta_i)| \leq \delta(\ell_{i+1} - \ell_i)/4$ for all $i \in N$.

2. $|\Theta_i| \geq 3$ for all $i \in N$, and for every $i \in N$ and every $\theta^j \in \Theta_i$, we have $v_i(\theta_i^{i+1}) - v_i(\theta_i^j) \leq \frac{\delta c^*}{8}$.

Then, $OPT(DecLin) = OPT(Cent)$.

The proof of Theorem 4.1 can be found in Appendix B. To better understand the conditions in the theorem, we briefly discuss the intuition behind them. The first condition implies the existence of an ‘interior solution’ in which all $n$ agents are active (in the centralized optimal solution). This is automatically satisfied if there is a profile of costs for which the virtual costs of all firms coincide. For example, this will immediately be the case if all firms share the same cost distribution. The second condition essentially requires the difference in the virtual costs between adjacent points in the support to be bounded by a function of $\delta$; the smaller the $\delta$, the closer the virtual costs should be. In general, if we think of the discrete distribution as an approximation of an underlying continuous distribution, this condition is equivalent to requiring that the grid of points in the support is thin enough with respect to $\delta$. For example, if costs are uniformly distributed in $[0,1]$, we can construct a grid consisting of $k$ equidistant costs such that the distance between adjacent costs is $1/(k-1)$. Using the definition of virtual costs (Definition 3.1), it is easy to see that the difference between adjacent virtual costs $v_i(\theta_i^{i+1}) - v_i(\theta_i^j)$ is bounded by $2/(k-1)$. Therefore, for every $\delta$, we can define $k$ large enough so that Condition 2 is satisfied (e.g. $k \geq 16/\delta c^*$).

Based on this discussion, we believe the conditions imposed in Theorem 4.1 are not too restrictive. In fact, both conditions will be satisfied provided that for at least one cost realization firms have similar virtual costs, and that the cost distribution grids are granular enough. These conditions together imply the existence of enough inter-related price vectors, providing sufficient degrees of freedom to satisfy the demand and the interim expected transfers constraints simultaneously.

4.4 Optimal Linear Pricing Decentralized Mechanism for Affine Demand

We now turn our attention to the more general affine demand models introduced in Section 2, and again ask when we can characterize the solution to the decentralized problem. By Corollary 4.1, the above intuition of the centralized optimal allocation holds for the optimal solution to the decentralized problem as well, whenever the optima of the two problems coincide. We next show
that (under sufficient mild conditions) we can guarantee $\text{OPT}(\text{DecLin}) = \text{OPT}(\text{Cent})$ by showing that the associated system of linear equations admits a solution.

**Theorem 4.2.** Consider the general setting with $N \geq 2$ agents, general costs distributions, and $\Gamma$ is strictly diagonally dominant. Suppose that the following conditions are simultaneously satisfied:

1. There exists a profile $\theta \in \Theta$ such that the set of active firms in the centralized optimal solution is $Q(\theta) = N$. In addition, there exists a $d^* \in \mathbb{R}$ such that, for all $\theta' \in \Theta$ with $|v_i(\theta_i') - v_i(\theta_i)| \leq d^*$ for all $i \in N$, we have $Q(\theta') = N$.

2. $|\Theta_i| \geq 3$ for all $i \in N$, and for every $i \in N$ and every $\theta_i^j \in \Theta_i$, we have $v_i(\theta_i^{j+1}) - v_i(\theta_i^j) \leq d^*/2$.

Then, $\text{OPT}(\text{DecLin}) = \text{OPT}(\text{Cent})$.

Although here $d^*$ depends on the primitives of the problem, the intuition behind the conditions is similar to that in the Hotelling model. The first condition implies the existence of a set of ‘interior solutions’ for which all firms are active. The second condition $d^*$ controls the distance between the virtual costs corresponding to adjacent costs. As we discussed after Theorem 4.1, this condition can always be satisfied by specifying a thin enough cost discretization, and the larger the set of interior solutions in (1), the coarser the discretization can be. The proof and the explicit characterization of $d^*$ for a few important classes of instances are deferred to the electronic companion.

**Incorporating Elastic Demand.** Throughout the paper we have assumed that total demand is inelastic: regardless of the prices, exactly one unit is consumed across all products in the assortment. While this assumption is natural in many contexts, we conclude this section by briefly discussing extensions that weaken this assumption. First, we consider the case where consumers still buy one unit across all products, but can also purchase from outside options; that is, for each product offered by a supplier, we assume that a product with the exact same characteristics can be obtained in the outside market at a fixed known price. In this case, all theorems extend almost straightforwardly. (The main difference is that the virtual costs of the products offered by supplier are now compared to both the virtual costs of other suppliers and the prices of the outside options.)

We also considered an extension in which we relax the constraint that demand should add up to one. Unfortunately, in this case it is in general not possible to find prices that simultaneously implement the centralized-optimal allocations and satisfy the constraints on interim expected transfers. In this case, we have less flexibility because to replicate these allocations prices need to be equal to virtual costs, as oppose to just matching virtual costs differences as in the inelastic demand model.
To gain a better understanding we numerically solved for the optimal decentralized mechanism in a set of instances. We found that, in general, suppliers with lower costs charge higher prices (thus serve less demand) in the decentralized problem than in the centralized one, even though the assortment of suppliers typically agreed in both problems. We believe these results are suggestive that the relaxation may give us an approximately optimal market structure that could serve as an input to simplify the solution to the decentralized problem. Further understanding optimal mechanisms under elastic demand seems like an interesting direction for future research.

5 First-Price Auction Implementation

In the previous section we characterized the optimal direct-revelation posted-price decentralized mechanism, which already captured a very important and realistic practical constraint: the use of posted linear prices without fixed upfront fees. The objective of this section is to move a step closer to practice by understanding how to incorporate a second important implementation constraint: in practice, FAs are implemented as first-price auctions (i.e., suppliers submit bids representing the unit prices of their products; if a product is added to the menu, the bid is taken as the posted price) with some additional rules to decide which products to include in the assortment. Unfortunately, one can prove that the optimal mechanism cannot generally be implemented using a first-price auction even for the demand systems considered in the paper; prices satisfying Eqs. (8) and (9) can be lower than suppliers’ costs for some realizations of cost vectors (while satisfying interim IR), but in a first-price auction no agent will bid lower than his cost. Therefore, in this section we rely on a combination of a simple theoretical model and numerical simulations to understand which are the main incentives issues arising when designing the rules to decide which products to include in the assortment, and how these rules can affect the performance of the auction.

The characterization of the linear pricing optimal mechanism is crucial for our purpose: it serves as a benchmark of what is achievable, and its structure also provides insights on how to modify the traditional first-price auctions to enhance performance. Specifically, the optimal mechanism provides insight on the optimal market structure, by the virtue of inheriting such properties from the optimal centralized solution. Also, recall that linear pricing imposes no optimality loss in the models studied in the paper. Therefore, in the context of this section, any losses observed in a first-price auction can be attributed to the particular rules of the auction being implemented (and not to linear pricing), thus allowing us to directly suggest improvements to the rules under consideration. As we shall see in Section 6, the insights obtained in the present section played a crucial role in the
Competition For the Market and Competition In the Market  As previously mentioned, FAs are usually implemented as a first-price auction (FPA) with rules to decide which products to include in the menu; these rules are typically a function of the suppliers’ bids, the characteristics of the products offered, and characteristics of the demand side. In such auctions, there are two different (but possibly complementary) types of incentives for the suppliers to aggressively compete in prices. First, suppliers compete at the auction stage to become part of the assortment. Whether a supplier is included or not in the assortment depends on the rules of the auction and the bids; by placing a lower bid, a supplier typically increases his chances of being part of the assortment. We refer to the competition at the auction stage as competition for the market. However, even if a supplier is added to the assortment, he is not guaranteed any fixed amount of demand: once in the menu, there is competition between imperfect substitute products. Naturally, one would expect that by placing a lower bid, a supplier can (weakly) increase his market share. We refer to the competition for demand once in the menu as competition in the market. We highlight that the optimal mechanism imposes competition for the market by restricting the entry of some suppliers to the menu. There is also competition in the market, because firms in the menu split the demand according to the demand system. In the rest of the section, we study the effect these two types of competition have on bids and consumer surplus under different first-price auction designs.

5.1 Analytical Evaluation of Different Designs in a Simple Model

Following the auction theory tradition, we assume that firms have private costs and that, for a given mechanism, they play a pure strategy Bayesian Nash equilibrium (BNE). Unfortunately, deriving the bidding strategies analytically under general model primitives is challenging as profits are a function of all bids through the demand system. Moreover, analytically characterizing bidding strategies in FPAs when bidders are not symmetric is hard; except in the simpler settings as the one studied in this section, suppliers in our problem are typically asymmetric due to their product characteristics. Therefore, to derive analytical results we restrict our attention to a simple Hotelling model on the unit segment. We consider a problem with two i.i.d. potential sellers located at 0 and 1 respectively and with two possible cost realizations $\theta_L, \theta_H$. As before, $\delta$ is the transportation cost. The analysis of this simple model provides essential insights. Then, we test the robustness of these insights with numerical experiments in more general models. To obtain the reported results, we considered a wide range of model parameters with $\theta_L \in [10, 19], \theta_H \in [10.5, 20]$, probabilities
of low cost in $[0.05, 0.95]$, and $\delta \in [0.5, 15]$, and the own-price elasticities (of the low type in the optimal mechanism) varied in the range $[-8, -0.5]$. All proofs corresponding to statements in this section can be found in the electronic companion.

**No competition for the market (NC mechanism).** Perhaps the simplest auction design is one with no competition for the market: every supplier whose price does not exceed a reserve price is added to the menu and bids are taken as posted prices. However, suppliers still compete for demand inside the menu: the demand is split among the firms in the menu according to their bids and the demand model. This mechanism is equivalent to a pricing game with private costs and a reserve, and is an interesting benchmark by its own which will also be helpful when studying ChileCompra’s FAs. We analytically calculated the BNE pricing strategies of this game when the reserve price is equal to $\theta_H$. Using the equilibrium prices, we computed the expected consumer surplus (i.e. the negative of expected purchasing cost plus the overall transportation cost) and compared it to that of the optimal mechanism. Generally, the optimality gaps ranged between 2.5% and 18% for different parameters of this model.

In this simplified setting we say that the outcome of the mechanism is *single-award* if, whenever agents have different types, the low-cost agent obtains all the demand; otherwise, we say that the outcome of the mechanism is *split-award*. A key difference between the NC and the optimal mechanism is that split-award outcomes occur more frequently in the former. This difference is key to understand the optimality gaps: higher gaps are observed for the values of $\delta$ in which NC split awards and the optimal mechanism does not. Intuitively, when $\delta$ is close to zero, both mechanisms single-award and the gap is small —because consumers are highly price sensitive, competition in the market provides sufficient incentives for suppliers to price aggressively. In contrast, for large values of $\delta$, both mechanisms split-award: restricting entry is not profitable as consumers’ value is mostly derived from variety. Finally, for intermediate values of $\delta$, NC split awards and the optimal mechanism does not, and the highest optimality gaps are observed. In this regime, by not restricting entry, NC fails to obtain competitive bids for the low-types (relative to the optimum), which increases the purchasing costs and thus deteriorates the performance. This suggests that, in this regime, actively restricting entry by modifying the auction rules might lead to improvements.

**Introducing Competition For the Market.** We now show how simple changes to the rules of the NC auction can improve performance. The new auction rules will generate competition for the market by restricting the entry of inefficient suppliers to obtain lower bids, thus making the
single-award outcome more likely. This emulates the findings from the optimal mechanism, which restricts entry of suppliers to reduce their expected payments. However, introducing competition for the market might reduce prices but might also increase transportation costs (reduce variety), so restricting entry does not necessarily translates into a higher consumer surplus. This trade-off is analyzed in detail in the electronic companion; we discuss the main take-aways next.

We consider two possible changes in the auctions’ rules: restricting entry \textit{ex-ante} (before observing the bids) and restricting it \textit{ex-post} (as a function of the observed bids). First, suppose that we decide to restrict entry \textit{ex-ante}. Note that if one is able to optimize over the number of suppliers that will be admitted in the menu, the ex-ante mechanism will always outperform the NC mechanism, as the latter is an ex-ante mechanism in which all suppliers are added to the assortment. Therefore, in our simple model, we must understand when does choosing a single winner using a FPA outperforms the NC mechanism. We find that restricting entry ex-ante is more beneficial when the low cost outcome is more likely to occur and \( \delta \) is intermediate; the optimality gaps can be decreased by up to 30%. The main drawback of the ex-ante mechanism, however, is that it always chooses one supplier (or a fixed number of them) even when all have similar (or the same) bids. If two suppliers have similar bids, by adding both to the menu we obtain more variety at a similar purchasing cost, thus improving consumer surplus.

To understand the limitation of this lack of flexibility, we next study a class of mechanisms that restrict entry \textit{ex-post}, that is, for which the decision on whom to include in the menu is contingent on the bids submitted. This emulates more closely the optimal mechanism, in which the assortment is decided based on the reported costs. Using the intuition from the optimal mechanism, we propose the following parametric restricted-entry (RE) first-price mechanism. There is a reserve price \( R \) (which we assume equal to \( \theta_H \)) and a split parameter \( C \). If bids satisfy \(|b_1 - b_2| < C\), then both suppliers are added to the menu. If not, only the supplier with the lowest bid is included in the menu. If both suppliers are in the menu, they still compete in the market as before. Hence, the only difference with the NC mechanism is that we restrict the entry to the menu, and the split parameter \( C \) quantifies how restrictive the entry to the market is.

As we had already discussed, restricting entry may cause performance to be worse than NC’s, as single-award increases the transportation cost. We define the \textit{best restricted-entry mechanism} (BRE) by optimizing over the split-parameter \( C \) to maximize expected consumer surplus. Note that as \( C = \delta \) is always a possibility, the BRE cannot do worse than NC. In fact, whenever BRE improves over NC it must be by restricting entry. Consistent with our intuition, the regime in which the performance of BRE is the best compared to that of NC is for intermediate values of \( \delta \).
Differentiation cost

Expected total cost

Expected purchasing cost

BRE

NC

Optimal

This is illustrated in Figure 1, where the BRE mechanism restricts the entry whenever $\delta \leq 4.67$. By doing so, the assortments are similar to those obtained in the optimal mechanism, and the expected purchasing cost and consumer surplus are closer to the optimal one. When $\delta$ exceeds 4.67, the savings obtained in the purchases cannot compensate for the increase in transportation costs and, therefore, BRE and NC coincide beyond that point. The optimality gaps in the instances we analyzed were reduced by at least 20% (and usually more than 50%) for those combinations of parameters in which the optimal mechanism restricts the entry and Chilecompra does not. The largest optimality gap was reduced from 20% to 7%.

Overall, our analysis shows that, by emulating the optimal mechanism to make substitute products compete for the market, consumer surplus can be significantly increased relative to NC.
5.2 Robustness Results and More General Settings: Numerical Experiments

To extend our analysis and test the robustness of our insights, we numerically solved for the equilibria for NC, the ex-ante, and the BRE mechanism in more involved models, and compared their expected consumer surplus with that of the optimal mechanism. We replicate this simulation exercise for a wide range of environments, and summarize the main findings next.

The most important common finding is that, as suggested by the theory, restricting entry is highly beneficial in the cases in which the optimal mechanism restricts entry but the NC does not. When considering more general cost distributions, we find that the lack of competition in the NC has a higher impact when the distribution is left-skewed or normal-like, with optimality gaps typically above 10%, and as high as 25%. When using the BRE, optimality gaps typically decrease by at least 40% in the regime of interest, and the gap differences become smaller as $\delta$ increases. In addition, restricting entry with the BRE improves performance for a wider range of values of $\delta$ as the number of values in the support increases, as the auctioneer can use a more refined splitting rule. On the other hand, the ex-ante mechanism decreased the optimality gap by at least 15%, performing better when the distribution was left-skewed.

We also considered settings with up to seven suppliers. We find that the gap between the optimal and NC increases as the number of suppliers increases. In contrast, restricting entry performs better (with respect to the optimum) than in the two agent case; this gap was rarely more than 5%, and decreased as the number of suppliers increased. In general, by restricting entry we are able to decrease the optimality gap by an average of 25% (and often by more than 70%).

Finally, we considered the general affine demand model introduced in Section 2, which also includes vertical differentiation, and we varied the vertical qualities of the products and the own and cross-price elasticities. In this setting, introducing competition for the market also improves performance, but the benefits were smaller when the difference in qualities among products were higher. This is to be expected: both the simple ex-ante and BRE mechanisms ignore quality advantages. Therefore, they tend to be naturally biased towards the supplier with lowest quality provided that, as one would expect, the cost distribution of the high-quality supplier stochastically dominates that of the low-quality one. However, it is still remarkable that such simple mechanisms can achieve significant savings in a setting with vertical differentiation. In particular, in the BRE, we obtained an average improvement in the gap of 7% in the cases were the difference between the highest and lowest qualities was more than 20%; when products had similar qualities, the average improvement was 15%. In the ex-ante case, the gains are slightly lower (4% and 11% respectively).

To conclude, we highlight that, in both mechanisms, we used a unique parameter (number of
agents and $C$ respectively) to restrict the entry. Although we expect better results if richer mechanisms that incorporate product characteristics or quality differences are used, such mechanisms would also yield more complicated rules and thus, as we shall see in the next section, they might be harder to implement in practice.

**Summary of main insights.** To conclude, we summarize the main insights gained from this section. First, there are two main sources of competition in this market: competition in the market (which naturally arises when substitute products compete in the menu), and competition for the market (which must be enforced through the rules of the auction). When products are very close substitutes, there is no need to introduce competition for the market, as the competition in the market to increase demand ensures low prices. Similarly, when products are very far substitutes, introducing competition for the market is not beneficial as, even though it lowers prices, it damages variety. In the cases in between, we find that emulating the optimal mechanism to introduce competition for the market is highly beneficial. Finally, while introducing competition using simple and anonymous rules in settings where preferences are horizontal leads to big improvements, these benefits tend to be lower when there is more vertical differentiation as anonymous rules tend to introduce a bias towards low-quality suppliers. In the next section, we explain in detail how these findings were applied in the redesign of the Chilean FAs.

6 Case Study: Redesigning the Chilean Framework Agreements

So far we have focused on understanding how to design procurement mechanisms to create menus of products for heterogeneous consumers under important realistic practical constraints: linear pricing and first-price auction implementation. The emphasis on addressing these implementation requirements was motivated by our desire to understand how to design better FAs in practice.

During the past year, we have been actively collaborating with the Chilean government in redesigning auctions for FAs. As a case study, we next discuss the redesign of the forthcoming food FA by describing the old design and the changes we implemented in the new design. As discussed in the introduction, these changes were based primarily on the insights obtained in the current paper. The new design resulted from a collaboration between the authors, a research group in University of Chile, and Chilecompra. The design has already been approved by ChileCompra’s executive directors, it is now waiting approval from the central government auditing body, and it is expected to run in the second half of this year. Also, it is worth highlighting that, although the focus of this section is on the redesign of the food FA, the designs of many other of ChileCompra’s FAs,
including hardware, office supplies, cleaning products, and personal care products, suffer from the same inefficiencies described in the present section. To address this issue, we continue working with ChileCompra in understanding how to better design FAs for some of these other product categories.

The Current Design of the Food Framework Agreement. ChileCompra currently decides the assortment of products in a FA by running a FPA-type mechanism which works as follows. First, the types of products needed in the FA (e.g., cereal, pasta) are announced. Then, each supplier submits a bid for each item he intends to offer; an item stands for a completely specified product. For example, a box of Kellogg’s Corn Flakes containing 15oz. and one containing 17oz. are considered different items. Suppliers can bid for any item they want, as long as the type of these products are among those required by the government. For example, if “cereal” is among the types of products required, a bid for any type of cereal is allowed, regardless of the brand, size, etc. Bids are then evaluated using a scoring rule; all products whose scores are above a threshold are offered in the menu at the price specified by the supplier’s bid. In practice, scores are dominated by price, and for brevity we abstract away from the other features considered. Prices are compared only across identical items. As a result, the current FA implementation works as if running one first-price auction independently for each item offered by at least one supplier. Moreover, the price score for an item-supplier pair is assigned by comparing his price to the minimum price of an identical item. If a unique supplier is offering the item, he automatically obtains the maximum score regardless of the price. As the item definitions are narrow (only identical products are directly compared), in most cases there is a single supplier bidding for an item, resulting in him being added to the menu.

To illustrate, in the current FA for food products (public auction number 2239-20-LP09), a total of 8091 products were offered by 116 suppliers. Out of those items, 4549 were offered by a unique supplier who got the maximum price score for this item; as a result, all such items were added to the menu. Furthermore, even for items with at least two bidders, the data suggests that the current rules fail to generate competition for the market. In the food FA, there were over 23,000 bids and only 5% of these were rejected because bids (prices) were too high. Hence, given the current rules, bidders have hardly any incentives to aggressively compete for the market.

Based on these empirical observations, the current FA design works similarly to the NC mechanism described in the previous section. Recall that one of our insights is that, in many situations, the performance can be improved if thicker markets are created by making imperfect substitute products compete at the auction stage to be part of the menu. Therefore, at least in some product types, it might be beneficial to implement an auction that restricting the entry of close substi-
tute products. Our theoretical and numerical results convinced ChileCompra’s executive board members to re-design the FA auction based on this idea, as described next.

**Designing the Auction for the New Food FA.** The main challenge to implement our insights in practice is how to identify sets of products that should be directly competing at the auction. In our analysis, we assumed that the demand system was an input to our model and, based on this input, our results allowed us to identify how aggressively these products should be competing to become part of the assortment. In practice, estimating a full demand system is challenging and imposes strong data requirements. In particular, an initial exploration of the data suggests that a significant amount of price variation is driven by unobserved demand characteristics, introducing severe endogeneity issues without useful supply shifters directly available. Estimating a full demand system in such a context of public procurement is interesting by itself, and matter of an ongoing independent research project. For the redesign we took a simpler approach, defining groups of close-substitute products to compete in the same auction based on a close collaboration with the ChileCompra team, which included both aggregate data analysis and input from their experts.

To define these groups, we first divided the products required by the government into 60 main *categories* (e.g. pasta, rice, chicken meat). In turn, these categories where divided into *subcategories* (e.g. chicken legs, full chicken). Finally, products in each subcategory where divided by brand. This categorization plays a key role in the redesign, as it provides a natural way of grouping products. In fact, the auction to become part of the menu will be conducted at the subcategory level.

Using the new categorization, the first big change suggested was to standardize the units of measure. In the past, for example, a supplier offering 1 kg. of rice would not compete in the auction with another one offering 1.5 kg. of rice, even if they were the same exact products. In the new design, for every single product there is a unique standard unit of measure that all suppliers need to provide a price for. In the rice example, all suppliers must provide a price for 1 kg., and only for 1 kg. This change by itself should significantly increase competition at the auction stage, as suppliers can no longer use the size as a tool to differentiate their offerings. (Separate prices are required for products that are typically sold at bulk in large volumes.)

Further, in the new design, the auctions at the subcategory level can be in one of four buckets:

- **Very high competition**: all products in the subcategory participate in the same auction, even if they have different brands. Only products whose bid prices are within the lowest 20% of all bids are added to the assortment.

- **High competition**: an auction is held separately for each (product,brand) pair in the subcat-
egory. Only products whose bids are within the lowest 20% in their auction are added to the assortment. In this case, different brands in the same subcategory do not compete at the auction stage, because of perceived significant quality differences among brands.

- **Low competition:** only products whose bids are within the lowest 80% of the subcategory are added to the assortment. All products in the subcategory participate in the same auction regardless of their brand.

- **Very low competition:** an auction is held separately for each \(\text{product,brand}\) in the subcategory. Only products whose bids are within the lowest 80% in their auction are added to the assortment.

The motivation for the last two designs is two-fold: (1) In some subcategories the government wants to prioritize diversity and inclusiveness of suppliers; and (2) These designs will allow us to better estimate the impact of the proposed changes in the ‘high competition’ designs by acting as control groups. In addition, note that the ‘high competition’ and the ‘very low competition’ auctions will be held separately for each \(\text{product,brand}\) combination in the subcategory. This was partially motivated by our discussion in the previous section, where we showed that the benefits of making substitute products compete are diluted if these products have very different qualities: a rule accounting only for price differences will bias the auction in favor of the lowest-quality suppliers, so more sophisticated rules incorporating quality differences should be used. In practice, one could think of using rules that group product brands based on quality (e.g., dividing products into “high” and “low” quality) to make them compete at the auction stage. However, this classification might appear somewhat arbitrary for the outside observer even if it is rigorously based on data. As we are working in a public setting, we opted for the simplest design: whenever the difference in qualities among brands is evident, only products from identical brands compete in the same auction.

Note that the percentage of bidders that will be added to the assortment is fixed and not contingent on bids for all designs. That is, these rules are “ex-ante” even though “ex-post” mechanisms performed better in our analysis in the previous section and more closely resemble the optimal mechanism. This decision was mainly driven by the fact that the ex-ante rules appeared easier to understand, explain, and justify. Also, there was some concern that, under ex-post competition, some categories might end up with only one supplier (or few of them) even if there were many bidders, which might not be desirable. We view this as a first step and we are considering mechanisms that restrict competition ex-post for the re-designs of other FAs in the future.

In the new design, each of the 60 subcategories was assigned to one of the competition crite-
ria described above trying to balance different considerations. For example, the markets for some categories are naturally thin, hence, to make sure that there were enough suppliers in the assortment, these were assigned to the low or very low competition designs. We also tried to balance the assignment of subcategories to the high competition ‘treatments’ and low competition ‘controls’.

Overall, the new design incorporates several of the managerial insights derived in the previous sections. In particular, it increases competition significantly by standardizing units (even in the ‘very low competition’ auctions), and it further increases competition in the ‘high competition’ designs by restricting the number of suppliers that will be part of the assortment.

7 Conclusions and Extensions

We presented a model to study procurement mechanisms for differentiated products. We characterized the optimal mechanism under real-world practical constraints for a general class of affine demand models. Then, we used these results to understand how to better design first-price FA auctions in practice and, in particular, to redesign ChileCompra’s food FA.

Our basic model can be extended in several interesting directions. First, it may be interesting to study in more detail models with an elastic total demand. Second, even though in the paper we use affine demand models in order to obtain analytical solutions, it might be interesting to understand whether our framework can be applied to other demand systems such as parametric discrete choice models. In addition, the focus in this paper had been on introducing competition at the auction level. When thinking more broadly about online two-sided markets, an additional design lever available to intensify competition is the search algorithm selecting the products shown to consumers (see, e.g., Dinerstein et al. (2014) for empirical evidence). Exploring the connection with these settings is a fruitful avenue for future work.

References


32