Procurement Mechanisms for Differentiated Products*

Daniela Saban† Gabriel Y. Weintraub‡

September 2015

Abstract

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, offered by strategic suppliers. Heterogeneous consumers then buy their most preferred alternative from the assortment as needed. Framework agreements (FAs), widely used in the public sector, take this form; the central government runs the initial auction and then the public organizations (hospitals, schools, etc.) buy from the selected assortment. This type of mechanism is also relevant in other contexts, including private procurement settings and the design of drug formularies. When evaluating the bids, the procurement agency must consider the optimal trade-off between offering a richer menu of products for consumers versus offering less variety, hoping to engage the suppliers in a more aggressive price competition. We develop a mechanism design approach to study this problem. We characterize the optimal mechanism, which typically restricts the entry of close-substitute products to the assortment to induce more price competition among suppliers, without much damage to variety. We then use the optimal mechanism as a benchmark to evaluate the performance of the Chilean government procurement agency's current implementation of FAs, used to acquire US$2 billion worth of goods per year. Through a combination of theoretical and numerical results we show how the performance of such FAs can be considerably improved by introducing simple modifications to current practice which, similarly to the optimal mechanism, increase price competition among close substitutes.

*We would like to thank the executives from the Chilean government procurement agency, Dirección ChileCompra, and in particular Guillermo Burr and David Escobar for the insights and data shared with us. The opinions expressed here are those of the authors and do not necessarily reflect the positions of Dirección ChileCompra nor of its executives. We thank Omar Besbes, Yeon-Koo Che, Awi Federgruen, Yonatan Gur, Jacob Leshno, Marcelo Olivares, Sasa Pekec, Michael Riordan, Jay Sethuraman, Nicolas Stier, and seminar participants at various conferences and institutions for useful comments. This work was supported by the Chazen Institute of International Business and the Deming Center at Columbia Business School, as well as by a Chile-Columbia Fund grant.

†Graduate School of Business, Stanford University, Email: `dsaban@stanford.edu`

‡Columbia Business School, Email: `gweintraub@columbia.edu`

1 Introduction

In this paper we study the following class of procurement mechanisms. A set of consumers affiliated to a certain organization (e.g., a government, a university, or a private firm) need to buy certain type of product (e.g., computers) and have heterogeneous preferences over specific items within the category (e.g., over the different computer models). The organization has a central procurement agency which manages the procurement process in two steps. First, the agency runs a mechanism to select an assortment of suppliers offering different products at given unit prices. Then, the consumers can buy their most preferred alternative in the assortment at the agreed price as needs arise, without undergoing any additional tendering process.

This type of procurement mechanism is widely used in practice. Several private firms and universities use catalogs (or assortments) of 'selected' suppliers and products from which their workers or units can buy from.[1] In addition, 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 (Truong 2014).[2] Finally, our main motivating practical example are *framework agreements* (FAs), public procurement mechanisms used by governments worldwide. In a FA, the central government procurement agency selects an assortment of differentiated products through competitive bidding in an auction mechanism. Then, whenever a public organization needs to make a purchase (e.g., a school buying computers), it buys its most preferred product from the assortment. To illustrate the importance of FAs, in 2010 the European Union awarded €80 billion using FAs, accounting for 17% of the total value of all public procurement (European Commision 2012).

The rationale behind using these procurement mechanisms is to exploit the purchasing power of a large central buyer (e.g., the central government), while still providing the heterogeneous consumers or organizations with some flexibility to select the product that best adapts to their needs. Therefore, an important challenge for the procurement agency when running these mechanisms is how to account for the heterogeneous preferences of the consumers buying from the assortment. For example, in government FAs, while a public school may want to buy laptops with attractive graphics features, the department of treasury may need laptops with high processing power. In addition, some patients might find prosthesis of a certain brand to be more comfortable than those of a competing brand. Similarly, the organizations buying from the food FA might also have different needs such as dietary constraints (e.g., hospitals and environments with kids). In all these cases, the government has a direct interest in providing variety to its organizations. The main objective of this paper is to provide insights on how to achieve (some) variety in a cost efficient way.

In particular, this paper is one of the first in the literature to provide a formal economic analysis

[1] For example, Stanford University, University of Minnesota.
[2] For example, Aetna, Blue Shield of California, United Healthcare

of this type of procurement mechanisms, making three contributions: (1) we introduce a model for the problem faced by the procurement agency; (2) we characterize the optimal mechanism for constructing the assortments; and (3) we use these results to study the design of simpler mechanisms that are commonly used in practical settings of FAs. Since our main motivation is to improve our understanding of FAs, we focus on auction-type mechanisms because these are typically preferred in public procurement. However, our results also shed light on other buying mechanisms used in the real-world settings mentioned above. For example, the optimal mechanism we characterize provides a general benchmark to evaluate the performance of current practice. We describe our main contributions in more detail next.

Our first main contribution is introducing a model capturing the following fundamental trade-off faced by a procurement agency when buying differentiated products. On one hand, consumers buying from the assortment usually have heterogeneous preferences. Therefore, increasing product variety in the assortment may increase consumer satisfaction, as it becomes more likely that consumers will find a better product for their needs. On the other hand, reducing the number of products in the assortment may increase suppliers' incentives to bid aggressively so that their products have a better chance to be part of the small selection of items. Our model extends classic auction and mechanism design models to study this trade-off between product variety and payments to suppliers.

In our model, there is a set of risk-neutral suppliers offering differentiated products, which are imperfect substitutes of each other. In the tradition of the auctions literature, we assume that suppliers have private information about their costs. The central procurement agency (designer) uses an auction-type mechanism to determine a *menu*, that is, an assortment of differentiated products together with the unit prices. Then, consumers with private heterogeneous preferences buy their most preferred alternative in the menu, which induces aggregate demands over products. In the tradition of the assortment literature, we assume that the *aggregate demands* as functions of the assortment and prices are common knowledge, and are an input to the model. Given the demand model, the designer chooses a mechanism with the objective of maximizing expected consumer surplus; this objective captures both variety and price considerations.

Our second main contribution is the characterization of the optimal direct-revelation posted-price mechanism for a broad class of affine demand models. This class includes the classic horizontal Hotelling demand model and a pure vertical demand model as particular cases, as well as more general specifications with both horizontal and vertical sources of product differentiation. Affine demand models are commonly used in competition models (e.g. Vives (2001)) and we believe they provide a reasonable balance between tractability and generality in our setting. We find that the optimal mechanism typically restricts the entry of close-substitute products to the assortment by selecting only one or few products from that set; this induces more price competition among suppliers, without much damage to variety. The characterization of the optimal mechanism allows us to formally quantify this optimal trade-off between variety and prices in terms of suppliers' costs, product characteristics, and substitution patterns across products.

Relative to the traditional mechanism design problem, a distinctive feature of our formulation is that the auctioneer cannot directly decide how to allocate demand across the products. Instead, the auctioneer selects the menu and demands are then determined by the underlying preferences of the organizations. This difference introduces significant complexities in the analysis of the problem, and makes the analytical characterization of the optimal mechanisms harder to obtain. In addition, most of the previous work in auction and mechanism design assumes homogeneous products (with some notable exceptions discussed in Section 2). Our work advances these literatures by accounting for an endogenous demand system for differentiated products.

Our third main contribution is to improve our understanding of the performance of certain type of auction mechanisms used in practice. The optimal mechanisms previously characterized are rarely implemented due to their complexity. However, they serve as a powerful tool to study practical mechanisms: optimal mechanisms provide a benchmark on what is achievable, and their structure provide insights on how to improve current practice. We are particularly interested in the type of FAs run by our collaborator in this project, the Chilean government procurement agency (Dirección ChileCompra), which in 2013 bought US$2 billion worth of goods using FAs.[3]

An important observation that arises by looking at the data from ChileCompra's FAs is that, because product definitions are narrow and auctions for different products are run independently, there is a single supplier bidding and winning for many products. Hence, while these suppliers may compete for demand once in the assortment, there is little to none competition *for the market* (i.e., at the auction stage). We study whether the current FA performance can be improved by creating thicker markets, making imperfect substitute products compete to be in the menu. To this end, we provide an extensive theoretical and numerical analysis of the current implementation of ChileCompra's FAs. Then, using the insights gained from the optimal mechanism, we explore possible changes to ChileCompra's FA implementation with regards to the set of suppliers to include in the menu. We show how rules that restrict the entry of close-substitute products (generating competition for the market), can significantly decrease suppliers' bids, without much damage to the variety offered. Furthermore, we provide a detailed analysis that illustrates when it is profitable to restrict the entry as a function of the market primitives. Overall, our results show that simple modifications to current practice can result in a significant increase in performance in terms of

[3]This represented a 21% of the value of all public procurement in Chile; see Área de Estudios e Inteligencia de Negocios, Dirección ChileCompra (2014).

expected consumer surplus.

The rest of the paper is organized as follows. Section 2 describes related literature. In Section 3 we formulate the mechanism design problem faced by the designer. In Section 4, we describe the general solution approach that we use to solve for the optimal mechanism. In Section 5, we characterize the optimal mechanism for affine demand models. In Section 6, we discuss the design of practical mechanisms using ChileCompra as a case study. We conclude and discuss extensions in Section 7. The proofs of the main results are provided in the appendix at the end of the paper. All others proofs and additional technical material are deferred to a separate electronic companion.

2 Related literature

Our work is related to several streams of literature in economics and operations. As previously mentioned, our work extends classic work in mechanism design in the tradition of Myerson (1981) by considering an endogenous demands system; this difference adds significant challenges when solving for the optimal mechanism. Furthermore, in our problem the designer maximizes consumer surplus —as opposed to just minimizing payments to suppliers—, which also depends on the underlying preferences of consumers.

Our work is also related to the oligopoly pricing models that study the effects of entry and competition in consumer surplus (e.g., Tirole (1988)). The main difference is that, in our setting, the decision to enter the market is not freely made by firms. Instead, it is decided by the designer based on the information elicited in the auction. Further, in our setting there is asymmetric information about firms' costs.

In that sense, our work is more related to previous papers in procurement and regulation economics. For example, Dana and Spier (1994) studies 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, as opposed to our case in which unit prices are determined. Similarly, Anton and Gertler (2004) and McGuire and Riordan (1995) study the optimal mechanism with an endogenous market structure in a Hotelling model of product differentiation. However, unit prices are not part of the mechanism, and allocations are determined by the designer and not endogenously by a demand system like in our case. A general insight of this body of work is that the designer may single-source more frequently if firms have private cost information, to be able to exert more pressure on efficient suppliers to reveal their costs; this is similar to some of our insights.

Closer to our work, Wolinsky (1997) studies a spatial duopoly model where firms compete in both prices and quality. While the paper considers an endogenous demand, the analysis is restricted to solutions in which both firms have positive demands. Instead, we are particularly interested in solutions in which some firms may be left out of the assortment to induce more competition. In fact, in our model, the optimal assortment typically does not contain all suppliers.

Another stream of related work that considers endogenous market structures is that of split-award auctions or dual sourcing in economics and operations (Chaturvedi et al. 2014, Li and Debo 2009, Elmaghraby 2000, Riordan and Sappington 1989, Anton and Yao 1989). These papers do not 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 (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 of ChileCompra's FAs in Section 6 is closely related to the idea of using a Demsetz auction (Demsetz 1968) to introduce *competition for the market*. This section is also related to papers in group buying showing that committing to an exclusive purchase from 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 models of complete information with suppliers that share the same marginal costs. Our analysis extends theirs to an auction setting with asymmetric information regarding private costs.

Finally, to the best of our knowledge only two prior papers directly study framework agreements (FAs), which is one of the main objectives of our work. Albano and Sparro (2008) consider a Hotelling model of horizontal differentiation, in which firms are located equidistantly and the subset of potential suppliers with lowest bids are selected in the assortment. In our case, we consider a richer set of rules in which the assortment can depend on product characteristic or location. Further, their analysis assumes complete information about firms' costs. Gur et al. (2013) 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, and in which unit prices are determined by the mechanism.

3 Model and Problem Formulation

In this section, we present our model and a formulation of the auctioneer's problem as a mechanism design problem.

3.1 Model

We introduce a model of procurement mechanisms for differentiated products demand systems. The agents of the model are (*i*) an auctioneer (or designer); (*ii*) suppliers (or agents); and (*iii*) consumers. The designer runs an auction-type mechanism to construct a menu (i.e., an assortment of products with posted unit prices) based on the suppliers' offers. Then, consumers purchase their most preferred product from this menu at the agreed price. We describe the main elements of the model next.

3.1.1 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. To simplify the exposition, we initially assume that each supplier offers exactly one product. Hence, unless otherwise stated, firms and products share the same indexes. In Section 7 and in the electronic companion, we discuss the extension to the multi-product setting; it is worth highlighting that our main results hold under this extension. We assume suppliers are risk-neutral, so they seek to maximize expected profits.

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 Θ_i is a finite set of strictly positive real numbers. We index the elements of Θ_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 θ_i if his cost is θ_i. Let f_i be a probability mass function over Θ_i, where $f_i(\theta_i)$ represents the probability that supplier i is of type θ_i. Let $F_i(\theta_i^j) = \sum_{k \leq j} f_i(\theta_i^k)$ be the cumulative probability distribution. Let $\Theta = \Pi_i \Theta_i$ denote the type space.[4] Because suppliers' types are independent, the joint probability of $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$ is equal to $f(\boldsymbol{\theta}) = \Pi_{i=1}^n f_i(\theta_i)$. We denote the probability that all suppliers other than i have type $\boldsymbol{\theta}_{-i}$ by $f_{-i}(\boldsymbol{\theta}_{-i})$.[5]

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 pro-

[4] We use discrete type distributions for technical convenience, as explained in Section 5.1.

[5] We use boldfaces to denote vectors and matrices throughout the paper.

duction 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 framework agreements (FAs) represent only a small fraction of their total production (Gur et al. 2013).

3.1.2 Consumers

In the tradition of the assortment 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. We introduce the following assumption.

Assumption 3.1 (Demand system)**.** *Suppose that, from the set of potential suppliers N, we fix a subset $Q \subseteq N$ of suppliers to be in the assortment. Let $\boldsymbol{p}_Q = \{p_i\}_{i\in Q}$, be the vector of their unit prices. Then, for every set Q and vector $\boldsymbol{p}_Q$, we assume that the vector of demand functions is given by:*

$$\boldsymbol{d}(Q, \boldsymbol{p}_Q) = \{d_i(Q, \boldsymbol{p}_Q)\}_{i\in Q}, \tag{1}$$

where $d_i(Q, \boldsymbol{p}_Q)$ denotes the expected demand for product i under assortment Q and prices $\boldsymbol{p}_Q$, and is common knowledge. We assume $d_i(Q, \boldsymbol{p}_Q) = 0$ for $i \notin Q$. In addition, $\sum_{i\in Q} d_i(Q, \boldsymbol{p}_Q) = 1$, for all $Q \neq \emptyset$ and $\boldsymbol{p}_Q$.

Note that the demand system $\boldsymbol{d}(Q, \boldsymbol{p}_Q)$ depends on the prices and the characteristics of all the products in the assortment. We assume that total demand for products in the assortment is normalized to one, which essentially amounts to assuming that total demand is perfectly inelastic and that there does not exist an outside option. However, our results extend to the case in which each product (or a subset of them) is also offered by an outside supplier at a given price. We do not include this extension to simplify notation.

The assumption of a known demand system is plausible in the contexts discussed in the introduction, because a demand system can typically be estimated using available historical data or consumer surveys (Ackerberg et al. (2006)). We note that we assume the designer is able to predict aggregate demands for every fixed set of products and prices; however, preferences of a specific consumer may be private information.[6]

We assume that the auctioneer maximizes consumer surplus when solving for the optimal mechanism. Hence, we also need a consumer surplus function as an input to our model. Given a demand system, the study of the 'integrability problem' provides conditions under which the demand functions can be derived from the maximization of a single utility function (see, e.g., Mas-Colell et al.

[6] Our analysis in Sections 4 and 5 is also valid under the weaker assumption (relative to common knowledge) that the auctioneer, but not the bidders, knows the demand functions.

(1995) and Anderson et al. (1992)). For all demand systems that we consider in this paper, this utility function corresponds to the consumer surplus function. We formalize this in the following assumption. Let $CS(\boldsymbol{d}, \boldsymbol{p})$ be the consumer surplus for demand quantities $\boldsymbol{d}$ and prices $\boldsymbol{p}$.

Assumption 3.2 (Consumer Surplus)**.** *The expression for consumer surplus must satisfy for all* $\boldsymbol{p}$*:*

$$\begin{aligned}(d_1(N,\boldsymbol{p}),\ldots,d_n(N,\boldsymbol{p})) \quad \in \quad & \underset{\boldsymbol{x}}{\operatorname{argmax}} \; CS(\boldsymbol{x},\boldsymbol{p}) \; , \\ & s.t. \qquad \sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0 \quad \forall i \in N \; .\end{aligned} \tag{2}$$

In addition, we require that for all $i \in N$*, there exists a function* $K(\boldsymbol{d})$ *of the quantities demanded such that:*

$$CS(\boldsymbol{d},\boldsymbol{p}) \quad = \quad K(\boldsymbol{d}) - \sum_{i=1}^{n} p_i d_i \; , \tag{3}$$

that is, consumer surplus is quasi-linear.[7]

The assumption states that the quantities demanded given prices $\boldsymbol{p}$ when all products are part of the assortment (as defined in Assumption 3.1) maximize consumer surplus given those prices.[8] We emphasize that Assumption 3.2 holds for all demand models that are considered in the paper.

A natural way of micro-founding an aggregate demand system and an associated consumer surplus function is to start from a discrete choice model that describes individual consumption decisions. See Anderson et al. (1992) for a general discussion; Armstrong and Vickers (2014) also provide a more specific discussion for the affine demand models used below. To illustrate, we present a simple example of a Hotelling demand model of horizontal differentiation with two suppliers and linear 'transportation costs'.

Example 3.1 (Hotelling model with two suppliers)**.** *Consider the unit interval as the product space, with two potential suppliers located at the extremes of the interval. There is a continuum of consumers uniformly distributed on the product space. Each consumer demands one unit of good and incurs transportation costs that are linear in the distance between the consumer and the supplier. Consumer* j *located at* ℓ_j *derives the following utilities from consuming from the set of suppliers* $N = \{1, 2\}$*:*

$$u_{j1}(p_1) = -\left(\delta\ell_j + p_1\right) \qquad \textit{and} \qquad u_{j2}(p_2) = -\left(\delta(1-\ell_j) + p_2\right),$$

[7] The latter assumption is useful to solve the optimal mechanism design problem.

[8] Note that the solution of this maximization problem may set some of the demand quantities equal to zero.

where supplier 1 *(resp.* 2*) is assumed to be located at* 0 *(resp.* 1*) and* δ *is the transportation cost. As consumers are uniformly distributed on the* $[0,1]$ *segment, the aggregate demands can be derived from individual utilities as follows:*

$$d_1(N,\boldsymbol{p}) = \max\left\{0, \min\left\{1, \frac{p_2 - p_1 + \delta}{2\delta}\right\}\right\} \qquad \textit{and} \qquad d_2(N,\boldsymbol{p}) = \max\left\{0, \min\left\{1, \frac{p_1 - p_2 + \delta}{2\delta}\right\}\right\}$$

In addition, aggregating the individual utilities we can derive the expression for consumer surplus:

$$CS(\boldsymbol{d},\boldsymbol{p}) = -\left(\frac{\delta}{2}\left(d_1^2 + d_2^2\right) + p_1 d_1 + p_2 d_2\right),$$

where the first terms represent the transportation costs and the latter terms the monetary costs. Note that in this example $K(\boldsymbol{x}) = -\frac{\delta}{2}(d_1^2 + d_2^2)$*, which is equivalent to the total transportation cost incurred by those consumers buying from* i*.*

3.1.3 Auctioneer

The role of the auctioneer is to select or design an auction-type mechanism to construct the menu of products based on the suppliers' offers. As previously mentioned, the menu consists of a subset of suppliers and unit prices for their products. Once selected, the rules of the auction are common-knowledge.

The auctioneer is risk-neutral and her objective is to maximize expected consumer surplus; this objective incorporates both variety considerations and payments to suppliers. Note that achieving variety is a natural objective in many relevant contexts, such as those of government purchasing described in the introduction.[9] Further, we think that consumer surplus is an adequate objective in several of the applications we have in mind. First, one of the key objectives of public procurement agencies is to increase government savings through convenient prices.[10] Second, this is also the case in private procurement. Having said that, our framework and approach could also be extended to maximize social welfare instead, albeit with some non-trivial modifications.

3.2 Mechanism Design Problem Formulation

We provide a mechanism design formulation of the auctioneer's problem. We consider mechanisms implemented in Bayes Nash equilibria. By invoking the revelation principle, we restrict attention

[9] We note that in other contexts, it may not always be in the auctioneer's best interest to provide variety. For example, a government may not be interested in providing too many options regarding certain products such as soft drinks or ink pens. In these cases, our results provide a way of evaluating the cost of incorporating variety considering consumers' idiosyncratic preferences, when perhaps the designer prefers to offer one (or very few) products.

[10] See, e.g. "Plan Estrategico 2013-2015," Dirección ChileCompra.

to direct revelation mechanisms without loss of optimality. Hence, for given cost declarations, the designer selects a menu which consists of an assortment of products (or suppliers) and their unit prices. Formally, a 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 $\boldsymbol{\theta}$; and (b) the *price functions* $p_i : \Theta \to \mathbb{R}$, where $p_i(\boldsymbol{\theta})$ is the unit price for the item offered by supplier i when cost declarations are $\boldsymbol{\theta}$. Note that this formulation allows for multiple suppliers to be in the menu. We define $\boldsymbol{q} = (q_1, ..., q_n)$ and $\boldsymbol{p} = (p_1, ..., p_n)$. For given cost declarations $\boldsymbol{\theta}$, the menu is given by $(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{p}(\boldsymbol{\theta}))$. We also define the *allocation functions* $x_i : \Theta \to [0,1]$, where $x_i(\boldsymbol{\theta})$ is the quantity allocated to supplier i when cost declarations are $\boldsymbol{\theta}$. Let $\boldsymbol{x} = (x_1, \ldots, x_n)$.

For each realization of $\boldsymbol{\theta}$, given the menu $(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{p}(\boldsymbol{\theta}))$, consumer demand is determined by the underlying demand system. Hence, for given $(\boldsymbol{q}, \boldsymbol{p})$, the allocation function $\boldsymbol{x}$ is restricted by the demand constraints in Eq. (1). This is in sharp contrast with classic mechanism design theory, in which the designer specifies a payment (or transfer) function and an allocation function. In our case, the designer selects an assortment and unit prices and, given these, allocations are decided by consumers. As discussed below, these constraints on the allocations introduce significant additional complexities to solving the mechanism design problem.

In the optimal mechanism design problem, the designer maximizes its objective (in our case, expected consumer surplus) 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 θ_i and report θ_i' as:

$$U_i(\theta_i'|\theta_i) = \sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} f_{-i}(\boldsymbol{\theta}_{-i})\big(\big(p_i(\theta_i', \boldsymbol{\theta}_{-i}) - \theta_i\big)\, x_i(\theta_i', \boldsymbol{\theta}_{-i})\big), \tag{4}$$

where $\boldsymbol{\theta}_{-i}$ is the report of supplier i's competitors. In addition, the problem must also have constraints to ensure that the allocations are consistent with the underlying demand system (Demand). Using the above definitions, the auctioneer's optimal mechanism design problem can be formulated as follows:

$$
\begin{aligned}
[P_0] \qquad & \max_{\boldsymbol{q},\boldsymbol{p},\boldsymbol{x}} \quad \mathbb{E}_{\boldsymbol{\theta}}[CS(\boldsymbol{x}(\boldsymbol{\theta}), \boldsymbol{p}(\boldsymbol{\theta}))] \\
& \text{s.t.} \quad U_i(\theta_i|\theta_i) \geq U_i(\theta_i'|\theta_i) \qquad \forall i \in N,\ \forall \theta_i, \theta_i' \in \Theta_i && \text{(IC)} \\
& \qquad\quad U_i(\theta_i|\theta_i) \geq 0 \qquad \forall i \in N,\ \forall \theta_i \in \Theta_i && \text{(IR)} \\
& \qquad\quad \sum_{i \in N} x_i(\boldsymbol{\theta}) = 1 \qquad \forall \boldsymbol{\theta} \in \Theta, \qquad x_i(\boldsymbol{\theta}) \geq 0 \qquad \forall i \in N, \forall \boldsymbol{\theta} \in \Theta && \text{(Feas)} \\
& \qquad\quad x_i(\boldsymbol{\theta}) = d_i(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{p}(\boldsymbol{\theta})) \qquad \forall i \in N,\ \forall \boldsymbol{\theta} \in \Theta, && \text{(Demand)}
\end{aligned}
$$

where $d_i(\cdot)$ correspond to the demand system introduced in Assumption 3.1. Note that we abused notation to denote by $\boldsymbol{q}(\boldsymbol{\theta})$ the set of suppliers that are in the assortment given costs $\boldsymbol{\theta}$. In the next section we discuss our approach to solve the optimal mechanism design problem P_0.

4 General Solution Approach

Problem P_0 is a mixed integer mathematical program. Further, even if one relaxes the integrality of the variables $\boldsymbol{q}$, the program is typically non-convex because demand equations are often non-linear, even in simple cases (see Example 3.1). Our solution approach relies on relaxing these demand constraints and solving the *relaxed problem*. The advantage of doing this is that the relaxed problem admits an analytical solution, which can be obtained by extending standard mechanism design arguments based on the envelope theorem (Myerson 1981) adapted for the setting of discrete distributions (Vohra 2011).[11] Further, the relaxed optimal solution has an intuitive interpretation: it is how a central planner would allocate demands across different suppliers to maximize consumer surplus, if he could dictate how consumers should behave (i.e. determine aggregate demands). Note that the central planner we refer to here has the ability to allocate demand, but it does not have access to suppliers' private information and still needs to respect their individual rationality and incentive compatibility constraints.

Then, we provide conditions that guarantee the existence of unit prices $\boldsymbol{p}$ that are consistent with the optimal solution of the relaxed problem and satisfy the demand constraints. Specifically, we need to find prices that will satisfy the incentive compatibility and individual rationality constraints on the supplier side, and provide incentives to consumers so that they behave as the central planner would want them to —the aggregate demands under these prices will agree with the optimal allocations in the relaxed problem. If such prices $\boldsymbol{p}$ exist, the optimal solution to the relaxed

[11] Note that these arguments based on the envelope theorem cannot be directly applied to problem P_0, because of the presence of the demand constraints, that in principle, could affect which IC constraints bind.

problem can be achieved by the original problem P_0. In other words, we show the existence of prices that allow us to decentralize the solution to the relaxed (centralized) problem. We formalize this argument next.

First, we introduce a new set of variables $t_i : \Theta \to \mathbb{R}$, where $t_i(\boldsymbol{\theta}) = p_i(\boldsymbol{\theta})x_i(\boldsymbol{\theta})$ represents the total *transfer* (or payment) to supplier i for a given cost declaration $\boldsymbol{\theta}$. Relaxing the demand constraints from P_0 and noting that interim utilities (Eq. (4)) can be written in terms of total transfers $\boldsymbol{t}$ and allocations $\boldsymbol{x}$, we obtain the relaxed problem:

$$
\begin{aligned}
[P_1] \qquad & \max_{\boldsymbol{x},\boldsymbol{t}} \quad \mathbb{E}_{\boldsymbol{\theta}}\left[K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n} t_i(\boldsymbol{\theta})\right] \\
& \text{s.t.} \quad U_i(\theta_i|\theta_i) \geq U_i(\theta_i'|\theta_i) \qquad \forall i \in N,\ \forall \theta_i, \theta_i' \in \Theta_i \qquad \text{(IC)} \\
& \qquad\quad U_i(\theta_i|\theta_i) \geq 0 \qquad \forall i \in N,\ \forall \theta_i \in \Theta_i \qquad \text{(IR)} \\
& \qquad\quad \sum_{i \in N} x_i(\boldsymbol{\theta}) = 1 \qquad \forall \boldsymbol{\theta} \in \Theta, \qquad x_i(\boldsymbol{\theta}) \geq 0 \qquad \forall i \in N,\ \boldsymbol{\theta} \in \Theta, \qquad \text{(Feas)}
\end{aligned}
$$

where $K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n} t_i(\boldsymbol{\theta})$ is the expression for consumer surplus given by Eq. (3) after replacing the second term (price times demand) by transfers.

We highlight that problem P_1 only differs from the classic mechanism design formulation in the objective function; while the traditional goal is to minimize expected transfers, we aim to maximize expected consumer surplus. Similarly 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 4.1 (Virtual costs)**.** *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 θ_i in Θ_i.*[12] *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 *of supplier i when he has type θ_i.*

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

Assumption 4.1 (Increasing virtual costs)**.** *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:

$$
\begin{aligned}
X_i(\theta_i) &\equiv \sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} f_{-i}(\boldsymbol{\theta}_{-i}) x_i(\theta_i, \boldsymbol{\theta}_{-i}), \\
T_i(\theta_i) &\equiv \sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} f_{-i}(\boldsymbol{\theta}_{-i}) t_i(\theta_i, \boldsymbol{\theta}_{-i}).
\end{aligned}
$$

[12] If θ_i is the lowest cost in the support, we define $\rho_i(\theta_i) = \theta_i$.

The advantage of solving the relaxed problem P_1 is that we can extend standard mechanism design arguments to characterize its optimal solution, as we formalize next.

Proposition 4.1. *Suppose that $(\boldsymbol{x}, \boldsymbol{t})$ satisfy the following conditions:*

1. *The allocation function satisfies for all $\boldsymbol{\theta} \in \Theta$,*

$$\begin{aligned} \boldsymbol{x}(\boldsymbol{\theta}) \quad &\in \quad \text{argmax} \;\; K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n} x_i(\boldsymbol{\theta}) v_i(\theta_i) \\ & \quad s.t. \qquad \sum_{i=1}^{n} x_i(\boldsymbol{\theta}) = 1, \qquad x_i(\boldsymbol{\theta}) \geq 0 \quad \forall i \in N \ . \end{aligned} \tag{5}$$

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 $\theta_i^j \in \Theta_i$:*

$$T_i(\theta_i^j) = \theta_i^j X_i(\theta_i^j) + \sum_{k=j+1}^{|\Theta_i|} (\theta_i^k - \theta_i^{k-1}) X_i(\theta_i^k) \tag{6}$$

Then, $(\boldsymbol{x}, \boldsymbol{t})$ is an optimal mechanism for problem P_1.

The proof can be found in Appendix A. Condition (1) in Proposition 4.1 states that, for each $\boldsymbol{\theta} \in \Theta$, the optimal vector of allocations $x(\boldsymbol{\theta})$ must be a maximizer of the consumer surplus when prices are set to be the virtual costs, subject to the feasibility constraints (see Eq. (3)). Further, by Eq. (2), the optimal solution is of the form $x_i(\boldsymbol{\theta}) = d_i(N, v(\boldsymbol{\theta}))$. Therefore, optimal allocations in P_1 have an intuitive form: they coincide with the demand functions given by Eq. (1) when the unit price of each supplier is exactly his virtual cost. This follows by the same arguments as in classic mechanism design, where 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.

It is important to note that, while the optimal demands are completely characterized, the optimal transfers are not. The only constraint imposed on transfers by the optimal solution is over interim expected transfers (Condition (3) in Proposition 4.1). As transfers are equal to unit price times demand, this implies that the optimal prices in the relaxed problem are underspecified. This freedom in the definition of optimal prices becomes useful later on, when we characterize the optimal solution to the original problem.

Because problem P_1 is a relaxation of P_0, 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 P_0 indeed attains the optimal objective of P_1.

Corollary 4.1. *Let* $(\boldsymbol{x}, \boldsymbol{T})$ *be the unique optimal solution to the relaxed problem* P_1.[13] *Define*

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

Suppose that for all $\boldsymbol{\theta} \in \Theta$*, there exist prices* $\boldsymbol{p}(\boldsymbol{\theta})$ *such that*

$$\begin{aligned} x_i(\boldsymbol{\theta}) &= d_i(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{p}(\boldsymbol{\theta})) \qquad \forall i \in N, \ \forall \boldsymbol{\theta} \in \Theta & (8) \\ & \quad \text{where } d_i(\boldsymbol{p}) \text{ is given by Eq. (1), and} \\ \sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} p_i(\theta_i, \boldsymbol{\theta}_{-i}) x_i(\theta_i, \boldsymbol{\theta}_{-i}) f_{-i}(\boldsymbol{\theta}_{-i}) &= T_i(\theta_i), \qquad \forall i \in N, \ \forall \theta_i \in \Theta_i \ . & (9) \end{aligned}$$

Then, the optimal objective of P_0 *is equal to the optimal objective of* P_1*. Moreover, an optimal solution of* P_0 *is given by* $(\boldsymbol{q}, \boldsymbol{p})$ *characterized by Eqs.* (7)*,* (8)*, and* (9)*, and the corresponding optimal allocation* $\boldsymbol{x}$ *of* P_0*. Furthermore, the optimal objective of* P_0 *is equal to the optimal objective of* P_1 *if and only if such solution* $(\boldsymbol{q}, \boldsymbol{p})$ *exists.*

The corollary suggests the following approach to solving the optimal mechanism design problem. First, solve the relaxed problem, the solution of which has an appealing structure —it gives us the solution a central-planner would choose to maximize consumer surplus. Then, find unit prices that allow to decentralize the optimal solution by making the aggregate demands under such prices agree with the relaxed optimal allocations, while satisfying the individual rationality and incentive compatibility constraints.[14] We use this solution approach in the next section to solve the original optimal mechanism design problem for different classes of affine demand models.

5 Affine Demand Models

The optimal mechanism design problem takes an underlying consumer demand model as an input. To obtain analytical solutions we will restrict attention to a general class of affine demand models, that is, models in which for every set Q, demands d as given by Eq. (1) are (piece-wise) affine functions of prices. Affine 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 (Allon and Federgruen 2007, Cachon and Harker 2002, Federgruen and Hu

[13]We denote $\boldsymbol{T}$ as the vector of interim expected transfers. Problem P_1 admits a unique optimal solution $(\boldsymbol{x}, \boldsymbol{T})$ for all demand systems considered in the paper. If P_1 admits more than one solution, our arguments can easily be extended accordingly.

[14]Here we differ from the topic of decentralizing efficient allocations in competitive equilibria (Mas-Colell et al. 1995), because we need to find prices that not only yield the desired allocations, but also provide suppliers' incentives for truthful revelation through the interim expected transfers.

2014). Furthermore, an additional advantage of these models in our setting is that they admit a convex and closed-form expression for consumer surplus.[15] It is easy to see that, even under affine demand models, the demand constraints are piece-wise linear, and problem P_0 remains non-convex.[16] However, the approach described above of relaxing these constraints will allow us to solve the problem.

In the remainder of this section, we first discuss how to apply the general solution approach introduced in Section 4 to affine demand models. Next, to gain intuition, we characterize the optimal mechanisms for a specific affine demand model: the classic Hotelling model of horizontal differentiation. Then, we provide the analysis for a general affine demand model that allows for products to be both horizontally and vertically differentiated.

5.1 Applying the Solution Approach to Affine Demand Models

We now discuss how to adapt the general solution approach described in Section 4 to affine demand models. Let $(\boldsymbol{x}, \boldsymbol{T})$ be the optimal solution to the relaxed problem P_1. By Proposition 4.1 and the discussion that follows, we have $x_i(\boldsymbol{\theta}) = d_i(N, \boldsymbol{v}(\boldsymbol{\theta}))$ where $\boldsymbol{v}(\boldsymbol{\theta})$ is defined as the vector of virtual costs, i.e., $\boldsymbol{v}(\boldsymbol{\theta}) = (v_1(\theta_1), \ldots, v_n(\theta_n))$. We denote $Q(\boldsymbol{\theta})$ as the set of *active suppliers* —those with strictly positive demands— in the optimal solution under cost realizations $\boldsymbol{\theta}$. By Corollary 4.1, to apply our solution approach we must find unit prices that simultaneously satisfy Eqs. (8) and (9).

Equations (8) require that unit prices $\boldsymbol{p}$ induce the optimal allocations $\boldsymbol{x}$ of P_1 through the demand system —as previously discussed, this is like decentralizing the allocations. Hence, we need to find unit prices such that $d_i(N, v(\boldsymbol{\theta})) = d_i(\boldsymbol{q}(\boldsymbol{\theta}), \boldsymbol{p}(\boldsymbol{\theta}))$ for all $i \in Q(\boldsymbol{\theta})$ and $\boldsymbol{\theta} \in \Theta$. 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. For each $\boldsymbol{\theta}$, these equations impose $|Q(\boldsymbol{\theta})|$ constraints over the prices $\boldsymbol{p}(\boldsymbol{\theta})$ corresponding to firms with strictly positive demands.[17] However, as the allocations must add

[15] An alternative to the class of affine demand models we use in this paper would be to start with a parametric discrete choice model, such as the multinomial logit model. These models are typically used in the assortment literature. Unfortunately, the basic multinomial logit model is not appropriate for our analysis because of its inability to capture substitution patterns due to the IIA property. An alternative that overcomes this issue is the multinomial logit model with random coefficients; however, this model is hard to solve even in the standard assortment problem, let alone in our auction setting. Another option that is typically more tractable in the traditional assortment problem is the nested logit model (Li and Rusmevichientong 2014). It may be worth studying in future work whether our framework can be applied to this demand system.

[16] For instance, consider the simple Hotelling model described in Example 3.1. There, the demand constraints for agent $i \in \{1, 2\}$ should be expressed as $x_i(\boldsymbol{\theta}) = \max\left\{0, \min\left\{1, \frac{p_j(\boldsymbol{\theta}) - p_i(\boldsymbol{\theta}) + \delta}{2\delta}\right\}\right\}$ with $j \in \{1, 2\}, j \neq i$, which yield a non-convex problem.

[17] In all demand models considered in the paper, only prices associated to suppliers with positive demand appear in the demand equations. This property is natural: if a supplier has zero demand, then its price does not play a role in the demand equations of competitors.

up to one, one of these constraints is redundant; the demands for $|Q(\boldsymbol{\theta})| - 1$ suppliers determines the demand for the remaining active supplier. Therefore, the equations in (8) impose $|Q(\boldsymbol{\theta})| - 1$ constraints over prices $\boldsymbol{p}(\boldsymbol{\theta})$. The redundancy of one constraint plays an important role because it induces degrees of freedom that can be used to satisfy the constraints on interim expected transfers.

In addition, Eqs. (9) require that unit prices $\boldsymbol{p}$ induce the interim expected transfers T_i in the optimal solution of P_1 —that is, the solution is individually rational, incentive compatible and interim expected payments to suppliers agree with those in the relaxed optimal solution. Given an optimal solution for P_1, $(\boldsymbol{x}, \boldsymbol{T})$, 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 P_1; therefore, the equations described in (9) are linear in unit prices. Also, observe that if in the optimal solution we have $x_i(\theta_i^j, \boldsymbol{\theta}_{-i}) = 0$ for all $\boldsymbol{\theta}_{-i} \in \Theta_{-i}$, then it must be that $T_i(\theta_i^j) = 0$. This follows by conditions (2) and (3) in Proposition 4.1. Hence, the previous equations impose $\sum_{i \in N} \sum_{\theta_i \in \Theta_i} \mathbb{I}[\exists\ \boldsymbol{\theta}_{-i} :\ i \in Q(\theta_i, \boldsymbol{\theta}_{-i})] \equiv K$ constraints ($\mathbb{I}[\cdot]$ denotes the indicator function).

By the observations above, verifying whether $OPT(P_0) = OPT(P_1)$ for affine demand models is equivalent to establishing whether the linear system of equations defined by Eqs. (8) and Eqs. (9) admits a solution. Here, $OPT(P)$ denotes the optimal value of problem P. Let $\boldsymbol{M}$ and $\boldsymbol{m}$ be the coefficient matrix and the corresponding right-hand side respectively defined by the linear equations in (8) and (9), where each column is associated with a price $p_i(\boldsymbol{\theta})$. We can safely discard the columns corresponding to prices $p_i(\boldsymbol{\theta})$ such that $i \notin Q(\boldsymbol{\theta})$, as all the coefficients of such columns are zero. The resulting matrix $\boldsymbol{M}$ will have $\sum_{\theta \in \Theta} |Q(\boldsymbol{\theta})|$ columns and $\sum_{\boldsymbol{\theta}} |Q(\boldsymbol{\theta})| - |\Theta| + K$ rows. It is easy to verify that $K \leq |\Theta|$ and, therefore, the number of columns is greater than or equal to the number of rows.

In the remainder of this section, we show that (under additional mild conditions) we can guarantee that the associated system of equations has a solution, and hence we can characterize the optimal mechanism.[18]

5.2 Optimal Mechanism for Hotelling Demand Model

Having described the general solution approach, we now discuss the structure of the optimal mechanism when the consumer demand is given by a general Hotelling model. This will allow us to provide intuition on the structure of the optimal mechanism, before discussing the general affine demand models in Section 5.3.

We now briefly discuss a general Hotelling demand model with an arbitrary number n of suppli-

[18] Assuming discrete types allow 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.

ers in the unitary segment. (Recall that a simple version of the Hotelling model was introduced in Example 3.1.) The n potential suppliers are located at $0 \leq \ell_1 < \ell_2 < \ldots < \ell_n \leq 1$ respectively; 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. The locations of the suppliers are assumed to be common-knowledge. 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 δ is the transportation cost and ℓ_j is the position of consumer j in the unit line.

Suppose that suppliers have fixed unit prices $\boldsymbol{p} = \{p_i\}_{i \in N}$. Then, the set of active suppliers with positive demand is given by $Q(\boldsymbol{p}) = \{i \in N : \ p_i \leq \min_{k \neq i} \{p_k + \delta|\ell_k - \ell_i|\}\}$, where we abused notation to make the set depend on prices instead of costs. In words, supplier i will be active if his price is lower than the total price (unit price plus transportation cost) a consumer at ℓ_i will pay if he buys from *any* other supplier. In this case, the consumers located in a neighborhood of ℓ_i choose to buy from supplier i. In addition, in the Hotelling model, suppliers split the market with their immediate active neighbors.[19] Note that the segment between two consecutive active suppliers i and j (that is, the segment between ℓ_i and ℓ_j) is divided proportionally to their prices: i obtains $\frac{p_j - p_i + \delta|\ell_j - \ell_i|}{2\delta}$ and j the rest.[20]

For the Hotelling model, the optimal solution to the relaxed problem P_1 is intuitive. By Proposition 4.1, the optimal allocations in the relaxed problem P_1 for a cost realization $\boldsymbol{\theta}$ are given by the Hotelling demands with prices equal to the vector of virtual costs $\boldsymbol{v}(\boldsymbol{\theta})$. Therefore, for a given vector of cost realizations $\boldsymbol{\theta}$, the optimal assortment is characterized as follows: $Q(\boldsymbol{\theta}) = \{i \in N : \ v_i(\theta_i) - v_j(\theta_j) \leq \delta|\ell_j - \ell_i| \quad \forall j \in N\}$, which corresponds to the definition of the set of active suppliers when prices are replaced by virtual costs.

To illustrate the result, consider Example 3.1 and suppose both suppliers have the same cost distribution. Let θ_1 and θ_2 be the cost realizations of supplier 1 and 2 respectively. In this case, the relaxed problem P_1 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.

An important feature of the relaxed optimal solution is that the decision of whether to split or not the demand depends on the cost realizations. In particular, if the transportation cost is small

[19] The demand equations can be easily derived by determining the location of an indifferent consumer between two active neighboring suppliers.

[20] If i is the leftmost active supplier, he obtains all the demand in the $[0, \ell_i]$ segment; similarly, if he is the rightmost active supplier, he obtains all the demand in $[\ell_i, 1]$.

relative to the differences in virtual costs, then the optimal solution includes only the supplier with the lowest virtual cost in the assortment. 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. Hence, by restricting the entry to the assortment in some scenarios, the auctioneer can reduce these expected payments while still providing incentives for truthful cost revelation. In contrast, if the transportation cost is large relative to the differences in virtual costs, the demand is split between the two suppliers to increase variety. In this sense, the optimal solution to the relaxed problem optimizes the trade-off between variety and payments to suppliers.

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[21]), then the optimal assortment will typically contain only the product with the lowest 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.

Optimal Solution to the Original Problem. We now describe the optimal solution to the original problem. By Corollary 4.1, if we can find a feasible pair $(\boldsymbol{q}, \boldsymbol{p})$ for P_0 such that the conditions of the corollary are satisfied, then we have found an optimal solution for the original problem; this solution will have exactly the same intuitive interpretation as the relaxed solution, because the assortment, allocations and expected payments agree. Therefore, we now study in which cases it is possible to achieve the same optimal objective in both the original problem and the relaxed problem, that is, in which cases $OPT(P_0) = OPT(P_1)$.

Consider the optimal solution of the relaxed problem as described by Proposition 4.1. Let q be defined as Corollary 4.1, that is, $q_i(\boldsymbol{\theta}) = 1$ if $i \in Q(\boldsymbol{\theta})$, and $q_i(\boldsymbol{\theta}) = 0$ otherwise, where $Q(\boldsymbol{\theta})$ is the set of active suppliers in the relaxed optimal solution under profile $\boldsymbol{\theta}$. Recall that in the Hotelling model, demands can be calculated based on the set of active suppliers; in particular, the segment between two consecutive active suppliers i and j is divided proportionally to their prices, where i obtains $\frac{p_j - p_i + \delta|\ell_j - \ell_i|}{2\delta}$ and j the rest. By comparing the Hotelling demands with the optimal allocations of P_1 as defined in Proposition 4.1, it should be clear that the constraints given by Eqs. (8) can be summarized as:

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

and $\vartheta_{\boldsymbol{\theta}}(i)$ denotes the successor of i in the set $Q(\boldsymbol{\theta})$ (i.e $\vartheta_{\boldsymbol{\theta}}(i) = \max j \in Q(\boldsymbol{\theta}) : \ j < i$). These

[21] Note that whether two products are close substitutes or not depends on both their relative distance in the product space, $|\ell_j - \ell_i|$, as well as how much relative weight consumers assign to product characteristics, summarized by the transportation cost δ.

constraints will implement the optimal allocations of P_1 using prices $\boldsymbol{p}(\boldsymbol{\theta})$. In words, the difference in prices between adjacent active suppliers must be equal to the difference in virtual costs. By Corollary 4.1, we must also guarantee that the expected transfers agree with the optimal ones, that is, unit prices should satisfy the constraints given by Eq. (9).

Unfortunately, even when demand is given by the Hotelling model, the optima of both problems might not always agree.[22] Therefore, we next focus on providing sufficient conditions under which $OPT(P_0) = OPT(P_1)$. The following theorem establishes that, under additional mild conditions, both optima will agree and thus the optimal mechanism can be characterized.

Theorem 5.1. *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* $\boldsymbol{\theta} \in \Theta$ *such that* $|v_{i+1}(\theta_{i+1}) - v_i(\theta_i)| \leq \delta(\ell_{i+1} - \ell_i)/4$ *for all* $i \in N$*; and*
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* $v_i(\theta_i^{j+1}) - v_i(\theta_i^j) \leq \frac{\delta c^*}{8}$.

Then, $OPT(P_0) = OPT(P_1)$.

The complete proof of Theorem 5.1 can be found in Appendix B.[23] We now briefly discuss the intuition behind the conditions. The second condition essentially requires the difference in the virtual costs between adjacent points in the support to be bounded by a function of δ; the smaller the δ, the closer the virtual costs should be. If we think of the discrete distribution as an approximation of an underlying continuous distribution, then this is equivalent to require the discretization to be thin enough with respect to δ. Intuitively, when the supports of the cost distributions are coarse, there are fewer combinations of prices and therefore fewer price vectors. As a result, there are not enough degrees of freedom to find prices that simultaneously satisfy the demand and the interim expected transfers constraints.

In addition, the first condition implies the existence of an 'interior solution' in which all n agents are active (in the relaxed optimal solution). We further require that there exist other solutions 'close' to that one —in which we replace the cost of an agent by one of his adjacent costs— where all agents are also active. Guaranteeing the existence of several cost profiles for which all agents are active translates into a structural relationship between the expected transfers constraints (Eq. (9))

[22] When suppliers have IID costs and are located at equidistant intervals, we show that the optima of the relaxed and the original problems always agree, by showing that the associated system of linear equations is consistent. In the electronic companion we provide an example which illustrates how the optima of both problems may fail to agree in the general case of non-IID costs.

[23] In Appendix B we prove a more general version of this theorem, which also accommodates the more general affine demand models defined in Section 5.3.

of all the agents. Intuitively, as the prices become more related with each other, there are more degrees of freedom to find prices that satisfy both the optimal demand constraints and the expected transfer constraints.

By the previous discussion, we believe the conditions imposed in Theorem 5.1 are not too restrictive. In fact, provided that the discretization of the support of the costs distributions is thin enough (relative to δ), and that for at least one cost realization all firms are active, both conditions will be satisfied.[24]

5.3 Optimal Mechanisms for General Affine Demand Models

We now study a more general affine demand model, that allow us to combine both vertical and horizontal sources of differentiation. Traditionally, an affine demand function is one where the relation $\boldsymbol{d}(\boldsymbol{p}) = \boldsymbol{\alpha} - \boldsymbol{\Gamma}\boldsymbol{p}$ holds for all $\boldsymbol{p} \in \{\boldsymbol{p} \in \mathbb{R} : \boldsymbol{\alpha} - \boldsymbol{\Gamma}\boldsymbol{p} \geq 0\}$. Here, $\boldsymbol{\alpha} \geq 0$ represents a quality (or vertical) component; $\boldsymbol{\Gamma}_{ij}$ represents the variation in the demand of product i as a result of a unit change in the price of product j, when all other prices remain constant. We assume that the products are substitutes, hence, $\boldsymbol{\Gamma}_{ij} \leq 0$ for $i \neq j$. For our purposes, it is important to consider the extension of this specification to price vectors under which some products get zero demand, as introduced by Shubik and Levitan (1980) and further analyzed by Soon et al. (2009).[25] We formalize this extension in our setting —where demands must add up to one—, by assuming that a single 'representative consumer' maximizes consumer surplus (Farahat and Perakis (2010) also use this approach to study oligopolistic pricing models under affine demand functions).[26]

We consider a representative consumer with a strictly concave gross utility function given by $u(\boldsymbol{x}) = \boldsymbol{c}'\boldsymbol{x} - \frac{1}{2}\boldsymbol{x}'\boldsymbol{D}\boldsymbol{x}$, where $\boldsymbol{D}$ is a positive definite matrix and $\boldsymbol{D}^{-1}$ is symmetric positive definite. The vector $\boldsymbol{c}'$ denotes the transpose of vector $\boldsymbol{c}$. Here, $\boldsymbol{D} = \boldsymbol{\Gamma}^{-1}$ and $\boldsymbol{c} = \boldsymbol{\Gamma}^{-1}\boldsymbol{\alpha}$ have been renamed to avoid burdensome notation. We further ask for $\boldsymbol{\Gamma}$ to be strictly diagonally dominant. The demand function is defined as the solution of the representative consumer's maximization problem, whose utility also corresponds to consumer surplus. That is, for any $\boldsymbol{p} \in \mathbb{R}^n$, let $\boldsymbol{d}(\boldsymbol{p})$ be defined as the solution of the following maximization problem:

[24] In the electronic companion, we provide a related characterization and result for a classic model of pure vertical differentiation. In this model products have different qualities on which all consumers agree upon; however, consumers have different price sensitives.

[25] Note that the Hotelling model presented in the previous section is a particular case of an affine demand model for which some products might get zero demand.

[26] Alternatively, a general affine demand model can also be micro-founded using consumers' individual utilities like in the Hotelling and vertical models (Martin 2009, Armstrong and Vickers 2014). However, we think the representative consumer approach provides a cleaner analysis.

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \boldsymbol{c}'\boldsymbol{x} - \frac{1}{2}\boldsymbol{x}'\boldsymbol{D}\boldsymbol{x} - \boldsymbol{p}'\boldsymbol{x} \\ s.t \quad & \boldsymbol{1}'\boldsymbol{x} = 1 \\ & \boldsymbol{x} \geq 0 \end{aligned} \tag{LD($\boldsymbol{p}$)}$$

Clearly, Problem (LD($\boldsymbol{p}$)) has a unique solution for every $\boldsymbol{p} \in \mathbb{R}^n$, and thus the demand function $\boldsymbol{d}(\boldsymbol{p})$ is well defined. To illustrate, we consider the following example:

Example 5.1. *We consider a duopoly where $\alpha = (a_1, a_2)$ and $\Gamma = \left(\begin{smallmatrix} r_1 & -\gamma \\ -\gamma & r_2 \end{smallmatrix}\right)$, with all the parameters positive and with $r_1 + r_2 \geq 2\gamma$. Under these parameters, we have $\boldsymbol{D} = \frac{1}{r_1 r_2 - \gamma^2}\left(\begin{smallmatrix} r_2 & \gamma \\ \gamma & r_1 \end{smallmatrix}\right)$ and $\boldsymbol{c} = \frac{1}{r_1 r_2 - \gamma^2}\left(\begin{smallmatrix} r_2 a_1 + \gamma a_2 \\ r_1 a_2 + \gamma a_1 \end{smallmatrix}\right)$. For any given $\boldsymbol{p}$, the demand functions $\boldsymbol{d}(\boldsymbol{p})$ are given by:*

$$d_1(\boldsymbol{p}) = \max\left\{0,\ \min\left\{\frac{(r_2-\gamma)a_1 - (r_1-\gamma)a_2 + r_1 - \gamma - (r_1 r_2 - \gamma^2)(p_1 - p_2)}{r_1 + r_2 - 2\gamma},\ 1\right\}\right\}$$

and

$$d_2(\boldsymbol{p}) = \max\left\{0,\ \min\left\{\frac{(r_1-\gamma)a_2 - (r_2-\gamma)a_1 + r_2 - \gamma - (r_1 r_2 - \gamma^2)(p_2 - p_1)}{r_1 + r_2 - 2\gamma},\ 1\right\}\right\}.$$

These demand functions exhibit natural properties; they are decreasing in a firm's own price and increasing in the competitor's price.[27] Also, depending on the price vector, there could be one or two firms active. In the appendix (Section B.1), we show that these and other properties hold for the general case with an arbitrary number of firms. In particular, we show that demands can be expressed as affine functions of prices of the set of active suppliers only, and that the demand for a product is weakly decreasing in his own price and increasing in others' prices. In addition, we show that demands only depend on price differences. This freedom in setting unit prices is essential to our proof technique as, similarly to the Hotelling case, we need to find unit prices that satisfy the same differences induced by the virtual costs, but that also simultaneously satisfy the interim expected transfer constraints.

As before, we note that the optimal allocations in the relaxed problem P_1 for a cost realization $\boldsymbol{\theta}$ are given by the demand characterization above with prices equal to the vector of virtual costs $\boldsymbol{v}(\boldsymbol{\theta})$. To illustrate, we discuss the structure of the optimal solution in Example 5.1. We focus, w.l.o.g. on supplier 1. For a given $\boldsymbol{\theta}$, he will be in the assortment ($d_1(\boldsymbol{\theta}) > 0$) if and only if $(r_1 r_2 - \gamma^2)(v_1(\boldsymbol{\theta}) - v_2(\boldsymbol{\theta})) \leq (r_2 - \gamma)a_1 - (r_1 - \gamma)a_2 + r_1 - \gamma$. Therefore, the difference in virtual costs for him to be active needs to be bounded by a quantity that is increasing in a sort of '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

[27] Note that the expressions are more complex relative to the initial model $\boldsymbol{d}(\boldsymbol{p}) = \boldsymbol{\alpha} - \boldsymbol{\Gamma}\boldsymbol{p}$, because of the constraint $\boldsymbol{1}'\boldsymbol{x} = 1$.

larger the difference in virtual costs allowed to keep supplier 1 active, which is intuitive. Note that similarly to the Hotelling model, the relaxed optimal solution may restrict the entry of a supplier to the assortment to decrease expected payments. The structure of the relaxed optimal allocation generalizes to the case of more products.

By Corollary 4.1, the above intuition of the relaxed optimal allocation holds for the optimal solution of the original problem as well, whenever the optima of the two problems coincide. Therefore, we next show that (under sufficient mild conditions) we can guarantee $OPT(P_0) = OPT(P_1)$, by showing that the associated system of linear equations admits a solution.

Theorem 5.2. *Consider the general setting in which $N \geq 2$ agents with general costs distributions. Suppose that the following conditions are simultaneously satisfied:*

1. *There exists a profile $\boldsymbol{\theta} \in \Theta$ such that the set of active firms in the relaxed optimal solution is $Q(\boldsymbol{\theta}) = N$. In addition, there exists a $d^* \in \mathbb{R}$ such that, for all $\boldsymbol{\theta}' \in \Theta$ with $|v_i(\theta_i') - v_i(\theta_i)| \leq d^*$ for all $i \in N$, we have $Q(\boldsymbol{\theta}') = N$.*
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, $OPT(P_0) = OPT(P_1)$.

We highlight that d^* depends on the primitives of the problem. In particular, we provide an explicit characterization of d^* for a class of instances in Appendix B.2. The intuition behind the conditions is similar to the Hotelling model; we must guarantee the existence of a set of 'interior solutions' and impose a 'thin enough' cost discretization.

6 Case Study: ChileCompra-Style Framework Agreements

In the previous section, we characterized the optimal directed-revelation posted-price mechanism. In practice, however, simpler mechanisms are generally used, as they are easier to explain to potential suppliers and require simpler management from the procurement agency. In particular, FAs are usually implemented as first price auctions with some additional rules to decide which products to include in the assortment. Unfortunately, one can prove that for all demand systems considered in the paper, in general the optimal mechanism *cannot* be implemented using a first price auction.[28] The objective of this section is to evaluate the performance of FAs that are used in practice and provide concrete recommendations for their improvement. The optimal mechanism

[28] One can show that, to be able to find prices that simultaneously satisfy Eqs. (8) and (9), for some some realizations of cost vectors the prices of some products might need to be lower than their actual costs (even though the expected interim utility is always positive by individual rationality). However, in a first price auction no agent will bid lower than his cost.

is crucial for this purpose: it serves as a benchmark of what is achievable, and its structure also provides insights on how to modify current practice to enhance performance.

The section is organized as follows. We start by describing the competition incentives that arise in first price auctions with additional rules to determine the assortment. In Section 6.2, we describe the FAs run by ChileCompra that will serve as our case-study. In Section 6.3, we use a simple model of horizontal differentiation to derive analytical results on the performance of ChileCompra-style FAs. Then, we quantify the potential improvements that can be achieved by introducing simple modifications to the current rules. Finally, in Section 6.4 we provide a large set of numerical experiments showing the robustness of the conclusions drawn from the analytical results in the simple model.

6.1 Competition For the Market and Competition In the Market

As previously mentioned, FAs are usually implemented as a first price auction (FPA) with some additional rules to decide which products to include in the assortment.[29] These rules are common-knowledge at the time of the auction, and are generally a function of the suppliers' bids, the characteristics of the products offered, as well as characteristics of the demand side. In such mechanisms, 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 (weakly) increases his chances of being part of the assortment. We refer to the competition at the auction stage as *competition for the market*.

Even if a supplier is added to the assortment, he is not guaranteed any fixed amount of demand. Once in the assortment, a supplier's final allocation depends on his own bid, the bids of the other suppliers in the assortment and the underlying demand system. Therefore, a supplier will be competing against imperfect substitute products for demand once in the assortment. 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 assortment 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 demand according to the demand system. In the rest of the section, we study the effect these two types of competition have on both the final bids (or prices) and the consumer surplus in ChileCompra-style FAs.

[29] By a first price auction we mean that suppliers submit bids, which represent the per-unit price of their products. If a product is added to the assortment, the bid is taken as the posted price.

6.2 ChileCompra's Framework Agreements

Since their introduction in 2004, FAs have been playing an increasingly important role in the procurement strategy of the Chilean government. In 2013, ChileCompra spent slightly more than US$ 2 billion in FAs, which corresponded to 21% of the total public expenditure in procurement and was twice the amount spent in 2010. Nowadays, more than 95 thousand products and services including food, office supplies, computers, and medical services can be acquired through FAs.

To award the FAs in a given category (e.g., food), ChileCompra runs a FPA-type mechanism which works as follows. First, ChileCompra announces the types of products needed within the category (e.g., cereal and pasta). Then, each supplier submits a bid for each *item* he intents 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 two 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 the brand, size, and so on.

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 in his bid. In practice, the scores are essentially dominated by price, so 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. Furthermore, the price score for an item-supplier pair is assigned by comparing his price to the minimum price of an identical item. If there is a unique supplier 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.

To illustrate, we consider the FA for food products.[30] There, 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 of 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. We highlight that many other FAs, such as office supplies, prosthesis, cleaning products, and personal care products, are similar to the FA for food products in that they create thin markets.

These observations motivate the following questions: can the performance of the current FAs

[30]This FA corresponds to the public auction number $2239-20-LP09$, titled "Alimentos Perecibles y No Perecibles", which was valid 2010 through 2014.

be improved if thicker markets are created by making imperfect substitute products compete to be in the menu? In other words, can competition *for the market*, in addition to competition *in the market*, improve performance?[31]

6.3 Analytical Evaluation of ChileCompra-Style FAs in 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). Hence, to evaluate the performance of the FA we need to derive such equilibrium bidding strategies. Unfortunately, deriving such strategies analytically under general model primitives is challenging as demands, and therefore profits, are a function of all bids through the demand system; to compute expected profits a bidder needs to integrate out over all possible demand realizations given competitors' bid functions.

Therefore, to be able to derive analytical results we restrict our attention to a simple pure horizontal differentiation Hotelling model. We consider a problem with two IID potential sellers located at 0 and 1 respectively in the unit line and with two cost realizations. Let $\Theta_i = \{\theta_L, \theta_H\}$ for $i = 1, 2$ and let f_L and f_H denote $f(\theta_L)$ and $f(\theta_H)$, respectively. As before, δ is the transportation cost. This simple model will provide essential insights. Then, we test the robustness of these insights with numerical experiments in more general models. All proofs in this section can be found in the electronic companion.

6.3.1 Analysis of ChileCompra-Style FAs

Supported both by the description of ChileCompra's mechanism and the analysis of their data, we propose the following first order approximation to their current FAs: we consider a procurement mechanism in which there is no competition to be in the menu, but suppliers must compete for demand inside the menu (i.e., there is no competition for the market but there is competition in the market). Every supplier whose price does not exceed the reserve price is added to the menu, and the bids of those suppliers are taken as posted prices. After, the demand is split among the agents in the menu according to the demand model.[32]

[31] Engel et al. (2002) also study this question on a stylized model of complete information. On related work, Klemperer (2010) proposes a 'product-mix' auction design for differentiated goods in the context of financial assets. Here, multiple 'varieties' are auctioned simultaneously, so that offers on one variety provide competitive pressure on the offers of other varieties. While that idea is similar to ours, the auctions and settings are different, and his paper does not provide mechanism design nor equilibrium results.

[32] A reserve price partially alleviates collusive concerns. As an example, if both suppliers increase their prices by the same amount in the Hotelling model, they obtain same allocations but a higher per-unit profit. The reserve price helps mitigating this collusive behavior. Designing mechanisms that deter collusion is certainly an important practical question, but it is out of the scope of the current paper.

We provide a theoretical analysis of the equilibrium bid functions and the performance of ChileCompra FA's in the two-by-two Hotelling model just described. We assume the ChileCompra mechanism imposes a reserve price equal to θ_H. Note that this mechanism is equivalent to a pricing game with private costs and a reserve. We analytically calculate the BNE strategies of this pricing game in the electronic companion. Using the equilibrium prices, we compute the expected consumer surplus (corresponding to the negative of expected supplier payments –purchasing cost– plus the overall transportation cost) of the ChileCompra mechanism and compare it to that of the optimal mechanism for different parameter values. To compare performance in this section, for a given a mechanism M, we define the *optimality gap* between the optimal mechanism and mechanism M as $(M/OPT - 1) * 100$, where we abuse notation and denote by M and OPT the total expected consumer surplus in mechanism M and the optimal mechanism, respectively. As an example, optimality gaps are shown in column 'Chile' in Table 1 as a function of both f_L and δ for fixed costs $\theta_L = 10$ and $\theta_H = 12$. Generally, the optimality gaps are between 2.5% and 18% for different model instances.[33]

In this simplified setting, we say that the outcome of the ChileCompra mechanism is *single-award* if, whenever agents have different types, the low-cost agent obtains all the demand when competing in the market. Otherwise, we say that the outcome of the mechanism is *split-award*.[34] A key difference between ChileCompra mechanism and the optimal mechanism is that the split-award outcome occurs more frequently in the former one; this difference helps understanding the optimality gaps. Higher gaps are observed for the values of δ in which ChileCompra split awards and the optimal mechanism does not. Intuitively, when δ is close to zero, both mechanisms single-award and the gap is small. In these cases, because consumers are highly price sensitive, competition in the market provides sufficient incentives for suppliers to price aggressively. In contrast, for large values of δ both mechanisms split-award; restricting entry is not profitable as consumers' value is mostly derived from variety. Finally, for intermediate values of δ, ChileCompra split awards and the optimal mechanism does not.

[33] We considered a wide range of model parameters with $\theta_L \in [10, 19]$, $\theta_H \in [10.5, 20]$, $f_L \in [0.05, 0.95]$, and $\delta \in [0.5, 15]$. For the model instances considered, the own-price elasticities (of the low type in the optimal mechanism) varied in the range $[-8, -0.5]$.

[34] We highlight that the terms single-award and split-award have been used in the literature with a different meaning. Typically, in auction settings, they refer to the outcome of the allocation rule (Anton and Yao 1989). In the context of the ChileCompra mechanism we use the terms to describe the outcome of competition. However, when referring to the optimal mechanism and the modification of the ChileCompra mechanism introduced below, we restore to the traditional meaning.

δ	$f_L = 0.1$		$f_L = 0.25$		$f_L = 0.5$		$f_L = 0.75$		$f_L = 0.9$	
	Chile	BRE	Chile	BRE	Chile	BRE	Chile	BRE	Chile	BRE
0.5	0.88	0.88	2.41	2.41	5.59	5.58	4.02	3.94	4.30	4.30
1	0.74	0.27	1.98	0.93	4.34	2.73	7.02	5.62	8.40	7.91
1.5	0.86	0.11	2.36	0.55	5.45	2.23	9.41	5.36	12.18	8.55
2	0.89	0.10	2.55	0.55	6.36	2.38	11.62	6.04	15.70	9.97
2.5	0.71	0.12	2.15	0.64	5.74	2.72	11.12	6.93	15.40	15.40
3	0.58	0.18	1.77	0.74	5.15	2.99	10.55	7.55	14.99	14.99
3.5	0.50	0.28	1.50	0.88	4.57	3.12	10.01	7.93	14.60	14.60
4	0.43	0.40	1.30	1.09	4.00	3.20	9.47	8.05	14.24	14.24
4.5	0.38	0.38	1.14	1.14	3.50	3.27	8.95	8.17	13.86	13.86
5	0.34	0.34	1.02	1.02	3.11	3.11	8.44	8.06	13.49	13.49
5.5	0.30	0.30	0.91	0.91	2.79	2.79	7.94	7.94	13.13	13.13
6	0.27	0.27	0.83	0.83	2.53	2.53	7.46	7.46	12.78	12.78

Table 1: Optimality gaps as a function of both the differentiation cost δ and f_L. The parameters are $\theta_L = 10, \theta_H = 12$. The horizontal lines indicate the point up to which restricting the entry outperforms ChileCompra's policy.

6.3.2 Analysis of Mechanisms that Introduce Competition For the Market

Recall that the optimal mechanism restricts entry of suppliers to reduce their expected payments. We now explore how the introduction of simple changes to the rules of ChileCompra's mechanism, which generate competition for the market to emulate the optimal mechanism, improves performance. The new auction rules make the single-award outcome more likely, restricting the entry of inefficient suppliers, obtaining lower bids.

Following ChileCompra's original design, we focus on FPA-type of mechanisms. We consider two possible changes in the auctions' rules: restricting entry ex-ante –before observing the bids– and restricting ex-post –as a function of the observed bids–.

Ex-Ante Restricted-Entry Mechanism. We start by analyzing what happens if competition for the market is induced by restricting entry before bids are placed. In particular, suppose that we decide how many agents will be in the menu before observing the bids and then run a FPA type mechanism to decide the prices. In our simple model, this amounts to deciding when does choosing a single winner using a FPA outperforms ChileCompra's mechanism in which all firms compete in the market (but the highest cost is sometimes priced out). A detailed analysis is provided in the electronic companion, but we now discuss the main take-away.

We observe that the simple modification to the FA rules can sometimes improve performance over the current mechanism, as competition for the market reduces prices. However, there is still a large set of parameters for which this is not the case. The main drawback of this type of mechanisms is that they always choose one supplier (or a fixed number of them) even if they have similar (or the same) bids. If two suppliers have similar bids, by adding both to the menu we obtain more variety (decrease transportation cost) at a similar purchasing cost, thus improving consumer surplus. This lack of flexibility is what damages the performance of mechanisms in which entry is restricted ex-ante. We discuss the performance of more sophisticated mechanisms next.

Ex-Post Restricted-Entry Mechanism. The main issue with restricting entry ex-ante is that such mechanisms do not split-award when suppliers submit the same bid, which causes an increase in the transportation cost. Therefore, we now study a class of mechanisms for which the decision on whom will be in the menu is contingent on the bids received by the auctioneer. Note that this emulates more closely the optimal mechanism; in the latter the assortment decisions are made as a function of 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 θ_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 lowest bid supplier (provided the bid is smaller than R) is included in the menu. Note that if both suppliers are in the menu they will still compete in the market as before. Hence, the only difference with ChileCompra's mechanism is that we restrict the entry to the menu and the split parameter C quantifies how restrictive the entry to the market is. Note that whenever $C = \delta$, our mechanism coincides with ChileCompra's, because in this simple Hotelling model the two suppliers are active if and only if the differences in prices is lower than δ.[35]

For the set of parameters in which ChileCompra single-awards, it can be shown that the performance of the mechanism cannot be improved by restricting entry. Therefore, our focus is in the settings in which ChileCompra split-awards. For these cases, we find values of C (smaller than δ) for which the equilibrium bid of the low-type induces single-award. Note that the equilibrium bid of a high-type is θ_H and, hence, a natural candidate for low-type equilibrium bid is $\theta_H - C$, because it is the highest possible bid that results in single-award. We have the following result.

Proposition 6.1. *For every set of parameters f_L, θ_H, θ_L and δ, there exists a (possibly empty) interval $\mathcal{I}$ such that, for all $C \in \mathcal{I}$, we have that $\theta_H - C$ is the unique equilibrium bid for the low type in the RE mechanism with reserve price θ_H and split parameter C.*[36]

For given model primitives, the designer is interested in maximizing expected consumer surplus. If restricting entry is a helpful device to achieve this objective, then the auctioneer will choose the

[35] Whenever $C = 0$, our mechanism agrees with a FPA. However, in this section we are only going to consider split parameters C for which a BNE exists; for discrete types a BNE may not exist for small values of C.

[36] In the electronic companion, we characterize the intervals referred to in the proposition as a function of f_L, θ_H, θ_L, and δ.

largest C for which a single-award equilibrium exists, because that induces the lowest bid for the low type. Hence, we define the "*best low-type bid*" to be $\theta_H - C^*$, where C^* is the highest C for which $\theta_H - C$ is an equilibrium bid for the low-type. The characterization of the best low-type bids can be found in the electronic companion but we briefly discuss the intuition. Intuitively, the advantage of bidding at $\theta_H - C$ is to capture the whole demand when the other agent has a high cost. As f_L becomes close to one, this advantage vanishes; for this reason, the best low-type bid is increasing in f_L. In addition, the best low-type bid is also increasing in δ; as the transportation cost increases, demands become less sensitive to prices and, therefore, a supplier can increase his bid without significantly decreasing demand.

Note that restricting entry may cause the performance to be worse than that of ChileCompra, as single-award increases the transportation cost. To that end, we define the *best restricted-entry mechanism* (BRE) as the mechanism that maximizes consumer surplus. We obtain the following straightforward result.

Proposition 6.2. *For a given set of parameters, the BRE has one of two possible forms: (1) coincides with the ChileCompra mechanism ($C = \delta$); or (2) uses the value C^* ($< \delta$) associated to the best low-type bid.*

For a given set of parameters, if BRE improves over ChileCompra it must be by restricting entry; in such case, (2) is optimal. Otherwise, (1) above is optimal. To illustrate, in Figure 1 we plot the outcome of the optimal, ChileCompra and BRE mechanisms as a function of the transportation cost δ for a given set parameters. As it can be observed, the BRE mechanism restricts the entry whenever $\delta \leq 4.675$. By doing so, the assortments obtained are similar to the ones generated by the optimal mechanism, and the expected purchasing cost is much closer to the optimal one. As a consequence, the expected total cost (or consumer surplus) becomes closer to the optimal one; in fact, for certain values of δ the optimality gap is reduced by more than half. When δ exceeds 4.675, the savings obtained in the purchases cannot compensate for the increase in transportation cost and, therefore, BRE and ChileCompra coincide beyond that point.

More generally, we study when BRE outperforms ChileCompra as a function of the parameters. We find that when δ is relatively small and ChileCompra split-awards, restricting entry improves over ChileCompra mechanism regardless of the value of other parameters. In such cases, the decrease in the low-type equilibrium bid results in a considerable decrease in the expected purchasing cost without a major increase in the expected transportation cost. In addition, as it can be observed in Table 1, restricting entry performs better for the middle-values of f_L. If f_L is too low, the savings are less likely to occur and therefore the potential impact is smaller. On the other hand, if f_L is too high, the best-low-type-bid tends to increase and the single-award becomes less profitable.

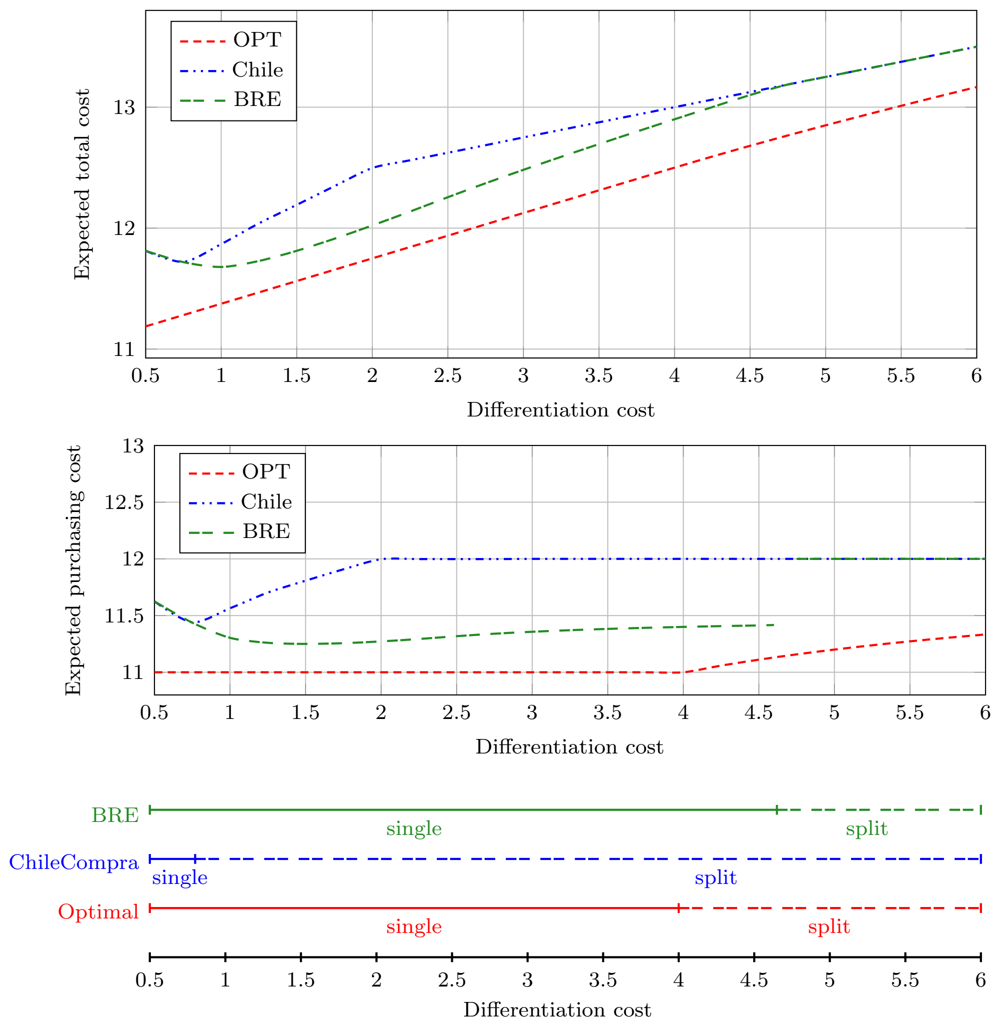

Figure 1: (*Top*) Expected total costs (purchasing plus transportation, equivalent to -(consumer surplus)) for optimal, ChileCompra and best restricted-entry (BRE) mechanisms as a function of the differentiation (transportation) cost δ. The parameters are $\theta_L = 10, \theta_H = 12$, $f_L = f_H = 1/2$. (*Center*) Expected purchasing costs. (*Bottom*) Single-award vs. split-award in optimal, ChileCompra's, and our BRE mechanism.

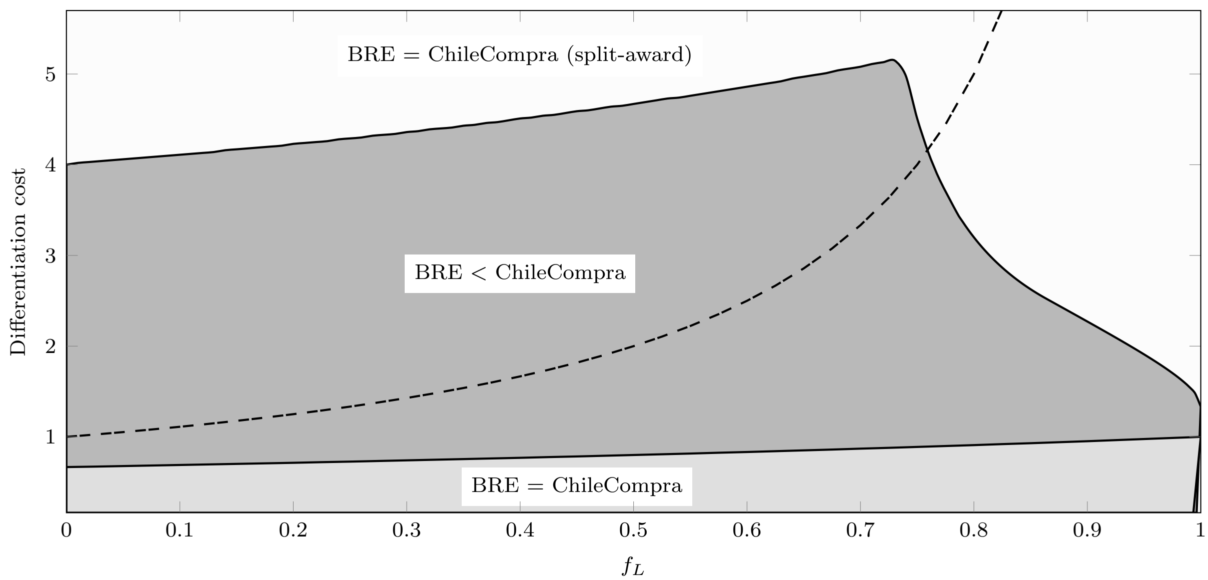

Figure 2: For $\theta_L = 10$, $\theta_H = 12$, we show when it is profitable to restrict the entry as a function of the differentiation cost δ and f_L. The dashed line represents the cutoff between single and split award in the optimal mechanism (i.e., $\delta = \frac{1}{f_H}(\theta_H - \theta_L)$).

This is illustrated by Figure 2, where we fix $\theta_L = 10$, $\theta_H = 12$, and show when it is profitable to restrict entry as a function of δ and f_L.

Overall, our analysis shows that ChileCompra FAs induce thin markets and that, by emulating the optimal mechanism to make substitute products compete to be in the assortment, consumer surplus can be significantly increased. Specifically, the optimality gaps in our model instances are 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 (e.g., see Table 1). The largest optimality gap was reduced from 20% to 7%.

We conclude this section with a note on the practical implementability of the restricted entry mechanisms. The BRE mechanism uses the best split-parameter C that depends on the problem primitives and therefore it may be hard to estimate in practice. However, we argue that even implementing the BRE mechanism with a rough estimate of the best C (but not the exact one) typically improves performance.[37] In particular, if restricting entry is profitable, any smaller C which is relatively close to the best C will induce the equilibrium bid $\theta_H - C$. Therefore, if the parameters are in the interior of the gray area in Figure 2 (where restricting entry improves performance), by choosing a conservative C the auctioneer should be able to increase consumer surplus. Also note that any C larger than the best C yields the same outcome as the current ChileCompra mechanism, so it will not damage performance.

[37] This is formally shown in the electronic companion.

6.4 Robustness Results: Numerical Experiments

To test the robustness of our intuition, we numerically solve for the equilibrium strategies for ChileCompra and the BRE mechanism and compare the expected consumer surplus in these mechanisms with that of the optimal. We replicate this simulation exercise for a range of environments by varying the cost distributions, the number of bidders and the demand model. The results are summarized next.

More General Cost Distributions. We first consider adding more points to the support of the cost distributions. To that end, we consider an initial interval and discretize it evenly into k costs, for $k = 2, 3, 5, 7$. We considered 4 types of distributions: uniform, left-skewed, right-skewed, and symmetric-unimodular (normal-like). We considered multiple combinations with $\theta_L \in [10, 19]$, $\theta_H \in [10.5, 20]$; here, θ_H and θ_L represent the maximum and minimum values in the support respectively. For each of those combinations, we vary $\delta \in [0.5, 15]$.[38] We highlight that, even though now we have multiple costs in the support, the auctioneer still must pick a unique split-parameter that remains fixed throughout the mechanism.

The results of our simulation show that the intuition for the multiple-costs case coincides with that of the two-by-two simple model and restricting entry improves the performance of the current mechanism. In general, the optimality gap decreases by at least 40%, and the differences in the gap become smaller as δ increases. Similarly to the two-by-two case, the relative benefits are greater when the distribution is left-skewed or normal-like where restricting entry achieves greater reduction in the bids of the low-type. In addition, as the number of values in the support increases, restricting entry improves performance for a wider range of values of cost-differentiation (δ), because the auctioneer can use a more refined splitting rule.

Larger Number of Bidders. We now consider models with more than two agents. To that end, we consider n agents at equidistant locations with agent i located at $\ell_i = (i-1)/(n-1)$. We test our results for $n \in \{2, 3, 4, 5\}$. The costs are still assumed to be IID across agents; however, agents are not ex-ante symmetric due to their locations. We consider the cost parameters (θ_L, θ_H) and f_L to vary in the same set as in the previous case ('More General Cost Distributions'). We allow δ to vary in the interval $[0.5, 30]$.

Whenever there are more than two agents, the auctioneer can choose whether to restrict entry either as a function of bids or as a joint function of both bids and product characteristics, as we discuss next. For the latter, the mechanism we consider is the restricted-entry mechanism with the

[38] As in the two-by-two case, in this section we combined the parameters so that the own-price elasticities where in the range $[-10, -0.5]$.

following modification: for split parameter C and bids $b_1, \ldots, b_n$, supplier i will be in the menu only if $b_i - b_j < C * |\ell_i - \ell_j|$ for every $j \in N$ with $j \neq i$. This rule is intuitive: it induces more price competition for agents that are close-by in the product space. Next, we consider restricting the entry solely as a function of bids. Similarly to the case of two agents, for split parameter C and bids $b_1, \ldots, b_n$, supplier i will be in the menu only if $b_i - b_j < C$ for every $j \in N$ with $j \neq i$. As this rule is less sophisticated than the previous one, a poorer performance is to be expected.

The main findings are as follows. First, the distribution of costs has the same impact in the performance as in the two-agent case; ChileCompra performs close to optimal for right-skewed distributions (when low-types rarely occur), but poorly for the other classes of distributions. Second, the optimality gap increases with the number of agents. The intuition seems to be the same as in the two-agent case; without competition for the market, ChileCompra fails to obtain competitive bids for the low-type (relative to the optimum) and this lack of competition has a higher impact as the number of suppliers increases. In accordance to what is observed in the two-agent case, restricting the entry improves performance for the sets of parameters in which ChileCompra split-awards. For the values of δ in which ChileCompra split-awards, restricting the entry performs better (with respect to the optimum) than in the two agent case. In general, the optimality gap for such instances decreases by an average of 29% (and in many cases by more than 80%) if characteristics are taken into account, and around 25% if they are not (in some instances, the gap is improved by more than 70%).

More General Demand Models. We now consider the general demand model introduced in Section 5.3, which includes both horizontal and vertical differentiation.

We first focus on the demand model in Example 5.1. We vary the qualities of the products and the own and cross price sensitivities. In particular, we consider the same set of costs and distributions as in the previous cases, and $r_1, r_2 \in [1, 4]$, $\gamma \in [0.05, 0.95]$ and $\alpha_1, \alpha_2 \in [12, 100]$, such that the combination of parameters yielded a positive expected consumer surplus under the optimal mechanism.

In this general model, suppliers are generally asymmetric ex-ante, as products can have different qualities. Similarly to the simplified two-agent case with horizontal differentiation, introducing competition for the market is more efficient whenever ChileCompra split-awards and the optimal mechanism single-awards. When quality differences between the two firms are large (more than 20% difference between α_1 and α_2), the optimality gaps decrease by an average of 8% when restricting entry. On the other hand, when quality differences are small, we obtain better results; a 22% decrease on average. This is to be expected; our simple restricted-entry mechanism only considers price differences, ignoring quality advantages. It is still remarkable though, that such a simple rule achieves significant savings in this setting with vertical differentiation.

Next, we consider the general demand model for more suppliers. Here, we only consider a single splitting parameter C, that is, supplier i is included in the menu if and only if $b_i - b_j \leq C$ for all $j \in N$. We again vary the qualities of the products and the own and cross price sensitivities. In particular, we consider the same set of costs and distributions as in the previous cases, $n \in \{2, 3, 4, 5\}$, $\gamma_{ii} \in [1, 10]$, $\gamma_{ij} \in [0.05, 0.95]$ and $\alpha_j \in [12, 300]$, such that the combination of parameters yielded a strictly diagonal dominant matrix $\mathbf{\Gamma}$ and a positive expected consumer surplus under the optimal mechanism. Consistently with what we observed in the two-agent setting, introducing competition for the market improves the performance in all cases, and it is more effective when products have similar qualities. Overall, restricting the entry allows us to decrease the optimality gap by an average of 10%. Similarly to the two-agent case, benefits were smaller when the difference in qualities were higher. We obtained an average improvement 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%. We highlight that we only used a unique parameter C to restrict the entry; we expect better results if multiple parameters C_{ij} are allowed (i.e. supplier i is in the assortment if and only if $b_i - b_j \leq C_{ij}$, for all j). We also expect better results for richer mechanisms that incorporate quality differences, for example, by favoring the entry of higher quality suppliers. However, such mechanisms also yield more complicated rules and thus may be harder to implement.

7 Conclusions and Extensions

In this paper we study procurement mechanisms for differentiated products demanded by heterogeneous consumers. First, we characterize the optimal mechanism for a general class of affine demand models. Second, we use these results to shed light on the FAs run by the Chilean government. Our results are useful to improve our understanding of FAs and, more generally, of buying mechanisms in similar contexts.

Our basic model can be extended in several interesting directions. First, to simplify the exposition, we assumed that each supplier offers one product. In the electronic companion, we provide an extension to our model in which we allow for multi-product suppliers. We show that our solution framework extends to this setting, and we are able to characterize the optimal mechanism for the multi-product case.

In addition, it would be interesting to further explore whether the insights are affected if other demand systems are assumed; for example a nested logit model and/or a model with an elastic downward sloping total demand function. Also, one might want to use econometric techniques to estimate important parameters of the model, such as those related to the underlying preferences of the organizations in the Chilean procurement setting, with the objective of sharpening the design recommendations. We leave all these directions for future research.

References

D. Ackerberg, C. Benkard, S. Berry, and A. Pakes. Econometric tools for analyzing market outcomes. *Handbook of Econometrics*, 2006.

Aetna. 2015 aetna pharmacy plan drug list information. URL `https://pbm.aetna.com/wps/portal/FRAMED_UNAUTH_MEDICATION_SEARCH_TOOL`.

G. L. Albano and M. Sparro. A simple model of framework agreements: Competition and efficiency. *Journal of Public Procurement*, 8(3):356–378, 2008.

G. Allon and A. Federgruen. Competition in service industries. *Operations Research*, 55(1):37–55, 2007.

S. P. Anderson, A. De Palma, and J. F. Thisse. *Discrete choice theory of product differentiation*. MIT press, 1992.

J. J. Anton and P. J. Gertler. Regulation, local monopolies and spatial competition. *Journal of Regulatory Economics*, 25(2):115–141, 2004.

J. J. Anton and D. A. Yao. Split awards, procurement, and innovation. *The RAND Journal of Economics*, pages 538–552, 1989.

Área de Estudios e Inteligencia de Negocios, Dirección ChileCompra. Informe de gestión, diciembre 2013. 2014.

M. Armstrong and J. Vickers. Which demand systems can be generated by discrete choice? Technical report, 2014.

Blue Shield of California. Drug formularies. URL `https://www.blueshieldca.com/bsca/pharmacy/formulary/home.sp`.

T. F. Bresnahan. Competition and collusion in the american automobile industry: The 1955 price war. *The Journal of Industrial Economics*, pages 457–482, 1987.

G. P. Cachon and P. T. Harker. Competition and outsourcing with scale economies. *Management Science*, 48(10):1314–1333, 2002.

D. Carlson and T. L. Markham. Schur complements of diagonally dominant matrices. *Czechoslovak Mathematical Journal*, 29(2):246–251, 1979.

A. Chaturvedi, D. R. Beil, and V. Martínez-de Albéniz. Split-award auctions for supplier retention. *Management Science*, 60(7):1719–1737, 2014.

Y. Chen and X. Li. Group buying commitment and sellers competitive advantages. *Journal of Economics & Management Strategy*, 22(1):164–183, 2013.

J. D. Dana. Buyer groups as strategic commitments. *Games and Economic Behavior*, 74(2):470–485, 2012.

J. J. Dana and K. E. Spier. Designing a private industry : Government auctions with endogenous market structure. *Journal of Public Economics*, 53(1):127–147, January 1994.

H. Demsetz. Why regulate utilities? *Journal of law and economics*, pages 55–65, 1968.

W. Elmaghraby. Supply contract competition and sourcing policies. *Manufacturing & Service Operations Management*, 2(4):350–371, 2000.

E. Engel, R. Fischer, and A. Galetovic. Competition in or for the field: Which is better? Working Paper 8869, National Bureau of Economic Research, April 2002.

European Commision. Annual public procurement implementation review, 2012.

A. Farahat and G. Perakis. A nonnegative extension of the affine demand function and equilibrium analysis for multiproduct price competition. *Operations Research Letters*, 38(4):280 – 286, 2010.

A. Federgruen and M. Hu. Sequential multi-product price competition in supply chain networks. Working paper, 2014.

Y. Gur, L. Lu, and G. Weintraub. Framework agreements in procurement: An auction model and design recommendation. Working paper, 2013.

R. G. Hansen. Auctions with endogenous quantity. *The RAND Journal of Economics*, pages 44–58, 1988.

R. A. Horn and C. R. Johnson. Topics in matrix analysis. *Cambridge University Presss, Cambridge*, 37:39, 1991.

P. Klemperer. The product-mix auction: a new auction design for differentiated goods. *ournal of the European Economic Association*, 8:526–536, 2010.

A. G. Kök, M. L. Fisher, and R. Vaidyanathan. Assortment planning: Review of literature and industry practice. In *Retail supply chain management*, pages 99–153. Springer, 2009.

V. Krishna. *Auction theory*. Academic press, 2009.

J. Levin. An optimal auction for complements. *Games and Economic Behavior*, 18(2):176–192, 1997.

C. Li and L. Debo. Second sourcing vs. sole sourcing with capacity investment and asymmetric information. *Manufacturing & Service Operations Management*, 11(3):448–470, 2009.

G. Li and P. Rusmevichientong. A greedy algorithm for the two-level nested logit model. *Operations Research Letters*, 42(5):319–324, 2014.

S. Martin. Microfoundations for the linear demand product differentiation model. Working Paper, 2009.

A. Mas-Colell, M. D. Whinston, J. R. Green, et al. *Microeconomic theory*, volume 1. Oxford university press New York, 1995.

T. G. McGuire and M. H. Riordan. Incomplete information and optimal market structure public purchases from private providers. *Journal of Public Economics*, 56(1):125–141, January 1995.

R. Myerson. Optimal auction design. *Mathematics of Operations Research*, 6(1):58–73, 1981.

M. H. Riordan and D. E. Sappington. Second sourcing. *The RAND Journal of Economics*, pages 41–58, 1989.

M. Shubik and R. Levitan. *Market Structure and Behavior*. Harvard University Press, Cambridge, MA, 1980.

W. Soon, G. Z. W., and J. Zhang. Complementarity demand functions and pricing models for multi-product markets. *European Journal of Applied Mathematics*, 20(5):399–430, 2009.

Stanford University. About iprocurement: Smartmart. URL http://web.stanford.edu/group/fms/fingate/finsystem/smartmart/index.html.

G. Strang. *Linear Algebra and Its Applications*. Brooks Cole, 3rd edition, 1988.

J. Tirole. *The Theory of Industrial Organization*. MIT press, 1988.

V.-A. Truong. Optimal selection of medical formularies. *Journal of Revenue & Pricing Management*, 13(2): 113–132, 2014.

United Healthcare. Drug formularies. URL https://www.uhcmedicaresolutions.com/health-plans/medicare-advantage-plans/resources-plan-material/drug-list.html.

University of Minnesota. U market. URL http://umarket.umn.edu.

X. Vives. *Oligopoly Pricing: Old Ideas and New Tools*. MIT Press, 2001.

R. Vohra. *Mechanism Design: A Linear Programming Approach*. Econometric Society Monographs. Cambridge University Press, 2011.

A. Wolinsky. Regulation of duopoly: managed competition vs regulated monopolies. *Journal of Economics & Management Strategy*, 6(4):821–847, 1997.

Main Appendix

A Proof of Proposition 4.1

Proof of Proposition 4.1. This proof uses the standard arguments from mechanism design theory introduced in Myerson's seminal paper (Myerson 1981). Since the supports of our cost distributions are discrete, we follow the version of these arguments presented by Vohra (2011). Throughout this proof, we define m_i to be the number of costs in the support of agent i, that is, $m_i = |\Theta_i|$.

We start by re-stating the IC and IR constraints in P_1 in terms of the expected allocations and transfers, as defined in Section 4:

$$
\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{t}} \quad & \mathbb{E}_{\boldsymbol{\theta}}\left[K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n} t_i(\boldsymbol{\theta})\right] \\
\text{s.t.} \quad & T_i(\theta_i) - X_i(\theta_i)\theta_i \geq T_i(\theta_i') - X_i(\theta_i')\theta_i \quad \forall i,\ \forall \theta_i, \theta_i' \in \Theta_i \\
& T_i(\theta_i) - X_i(\theta_i)\theta_i \geq 0 \quad \forall i,\ \forall \theta_i \in \Theta_i \\
& \sum_{i \in N} x_i(\boldsymbol{\theta}) = 1 \qquad \forall \boldsymbol{\theta} \in \Theta, \qquad x_i(\boldsymbol{\theta}) \geq 0 \qquad \forall i \in N,\ \boldsymbol{\theta} \in \Theta,
\end{aligned}
$$

Recall that $\Theta_i = \{\theta_i^1, ..., \theta_i^{m_i}\}$. If we add a dummy type per agent $\theta_i^{m_i+1}$ such that $X_i(\theta_i^{m_i+1}) = 0$ and $T_i(\theta_i^{m_i+1}) = 0$, then we can fold the IR constraints into the IC constraints:

$$
T_i(\theta_i^j) - X_i(\theta_i^j)\theta_i^j \geq T_i(\theta_i^k) - X_i(\theta_i^k)\theta_i^j \quad \forall j \in \{1, ..., m_i\},\ \forall k \in \{1, ..., m_{i+1}\}\ .
$$

Applying Theorem 6.2.1 in Vohra (2011) for our procurement setting we obtain that an allocation $\boldsymbol{x}$ is implementable in Bayes Nash equilibrium if and only if $X_i(\cdot)$ is monotonically decreasing for all $i = 1, ..., n$.[39] Further, by Theorem 6.2.2 in Vohra (2011), all IC constraints are implied by the following local IC constraints:

$$
\begin{cases}
T_i(\theta_i^j) - X_i(\theta_i^j)\theta_i^j \geq T_i(\theta_i^{j+1}) - X_i(\theta_i^{j+1})\theta_i^j & (BNIC_{i,\theta}^d) \\
T_i(\theta_i^j) - X_i(\theta_i^j)\theta_i^j \geq T_i(\theta_i^{j-1}) - X_i(\theta_i^{j-1})\theta_i^j & (BNIC_{i,\theta}^u)
\end{cases}
$$

[39] Note that the results cited in Vohra are for IID bidders, but the extension to bidders with different distributions is straightforward.

Therefore, we can re-write the problem as:

$$\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{t}} \quad & \mathbb{E}_{\boldsymbol{\theta}}\left[K(\boldsymbol{x}(\boldsymbol{\theta}))\right] - \sum_{i=1}^{n}\sum_{j=1}^{m_i} f_i(\theta_i^j)T_i(\theta_i^j) && (obj)\\
\text{s.t.} \quad & T_i(\theta_i^j) - X_i(\theta_i^j)\theta_i^j \geq T_i(\theta_i^{j+1}) - X_i(\theta_i^{j+1})\theta_i^j \quad \forall i \in N,\ \forall j \in \{1, ..., m_i\} && (BNIC_{i,j}^d)\\
& T_i(\theta_i^j) - X_i(\theta_i^j)\theta_i^j \geq T_i(\theta_i^{j-1}) - X_i(\theta_i^{j-1})\theta_i^j \quad \forall i \in N,\ \forall j \in \{2, ..., m_i\} && (BNIC_{i,j}^u)\\
& 0 \leq X_i(\theta^{m_i}) \leq \ldots \leq X_i(\theta^1), \quad \forall i \in N && \text{(M)}\\
& \sum_{i=1}^{n} x_i(\boldsymbol{\theta}) = 1 \qquad \forall \boldsymbol{\theta} \in \Theta, \qquad x_i(\boldsymbol{\theta}) \geq 0 \qquad \forall i \in N,\ \boldsymbol{\theta} \in \Theta.
\end{aligned}$$

Using standard arguments, we can show that all downward constraints $(BNIC_{i,j}^d)$ bind in the optimal solution.[40] Hence,

$$T_i(\theta_i^j) - X_i(\theta_i^j)\theta_i^j = T_i(\theta_i^{j+1}) - X_i(\theta_i^{j+1})\theta_i^j \quad \forall i \in N,\ \forall j \in \{1, ..., m_i\}.$$

Furthermore, it is simple to show that, in this case, the upward constraints $(BNIC_{i,j}^u)$ are satisfied. Applying the previous equation recursively we obtain:

$$T_i(\theta_i^j) = \theta_i^j X_i(\theta_i^j) + \sum_{k=j+1}^{m_i} (\theta^k - \theta^{k-1}) X_i(\theta_i^k) \ . \tag{11}$$

[40] A formal proof can be obtained by trivially adapting the Lemma 6.2.4 in Vohra to the procurement case.

Replacing in the objective:

$$\begin{aligned}
obj &= \mathbb{E}_{\boldsymbol{\theta}}\left[K(\boldsymbol{x}(\boldsymbol{\theta}))\right] - \sum_{i=1}^{n}\sum_{j=1}^{m_i} f_i(\theta_i^j)T_i(\theta_i^j) \\
&= \mathbb{E}_{\boldsymbol{\theta}}\left[K(\boldsymbol{x}(\boldsymbol{\theta}))\right] - \sum_{i=1}^{n}\sum_{j=1}^{m_i} f_i(\theta_i^j)\left(\theta_i^j X_i(\theta_i^j) + \sum_{k=j+1}^{m_i}(\theta_i^k - \theta_i^{k-1})X_i(\theta_i^k)\right) \\
&= \mathbb{E}_{\boldsymbol{\theta}}\left[K(\boldsymbol{x}(\boldsymbol{\theta}))\right] - \sum_{i=1}^{n}\sum_{j=1}^{m_i} f_i(\theta_i^j)\left(\theta^j X_i(\theta_i^j)\right) - \sum_{i=1}^{n}\sum_{j=1}^{m_i}\sum_{k=0}^{m_i-1} f_i(\theta_i^j)\left(\mathbb{I}\{k \geq j\}(\theta_i^{k+1} - \theta_i^k)X_i(\theta_i^{k+1})\right) \\
&= \mathbb{E}_{\boldsymbol{\theta}}\left[K(\boldsymbol{x}(\boldsymbol{\theta}))\right] - \sum_{i=1}^{n}\sum_{j=1}^{m_i} f_i(\theta_i^j)\left(\theta^j X_i(\theta_i^j)\right) - \sum_{i=1}^{n}\sum_{k=1}^{m_i} F_i(\theta_i^{k-1})(\theta_i^k - \theta_i^{k-1})X_i(\theta_i^k) \\
&= \sum_{\boldsymbol{\theta}\in\Theta} f(\boldsymbol{\theta})K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n}\sum_{j=1}^{m_i} f_i(\theta_i^j)\left(\left(\theta^j + \frac{F_i(\theta_i^{j-1})}{f_i(\theta_i^j)}(\theta_i^j - \theta_i^{j-1})\right)X_i(\theta_i^j)\right) \\
&= \sum_{\boldsymbol{\theta}\in\Theta} f(\boldsymbol{\theta})K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n}\sum_{\theta_i\in\Theta_i} f_i(\theta_i)v_i(\theta_i)X_i(\theta_i) \\
&= \sum_{\boldsymbol{\theta}\in\Theta} f(\boldsymbol{\theta})\left(K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n} v_i(\theta_i)x_i(\boldsymbol{\theta})\right)
\end{aligned}$$

The equations follow by simple algebra. In particular, the fourth equation follows by changing the order of summations. Therefore, if we find an allocation such that for all $\boldsymbol{\theta} \in \Theta$,

$$\begin{aligned}
x(\boldsymbol{\theta}) \quad &\in \quad \operatorname{argmax}\ K(\boldsymbol{x}(\boldsymbol{\theta})) - \sum_{i=1}^{n} x_i(\boldsymbol{\theta})v_i(\theta_i) \\
&\text{s.t.} \qquad \sum_{i=1}^{n} x_i(\boldsymbol{\theta}) = 1, \qquad x_i(\boldsymbol{\theta}) \geq 0 \quad \forall i \in N\ ;
\end{aligned}$$

such that the interim expected allocations are monotonic for all $i \in N$ (i.e. $X_i(\theta) \geq X_i(\theta')$ for all $\theta \leq \theta' \in \Theta_i$); and such that the interim expected transfers satisfy Eqs. (28), for all $i \in N$ and $\theta \in \Theta_i$, then we have found an optimal solution. $\square$

B Proof of Main Theorems

In this section we prove our main theorems. In particular, we prove a more general theorem (Theorem B.1), which generalizes the statements of Theorem 5.1 (for the Hotelling model) and Theorem 5.2 (for general affine demands).

The rest of the section is organized as follows. Recall that the idea of the proof is to show that the system of linear equations defined by Eqs. (8) and (9) is consistent (see Corollary 4.1). Therefore, in Section B.1 we start by describing the coefficient matrix of the associated system of equations, and deriving some properties of the matrix that will be useful to prove the theorem. In Section B.2, we state some definitions needed for our proof. In Section B.3, we state and prove a preliminary lemma that plays an important role in our proof. Finally, in Section B.4, we state and prove the main theorem. Naturally, we use several basic definitions and concepts from linear algebra throughout this section. We refer the reader to Strang (1988).

B.1 The Coefficient Matrix and the System of Equations

Given $\boldsymbol{\theta} \in \Theta$ and $i \in N$, let $\boldsymbol{A}_{ij}(\boldsymbol{\theta})$ denote the coefficient of $v_j(\theta_j)$ corresponding to the left hand side of Eqs. (8); that is, the coefficient of $v_j(\theta_j)$ in $d_i(N, v(\boldsymbol{\theta}))$. Recall that $Q(\boldsymbol{\theta})$ is the set of active firms in the relaxed optimal solution under profile $\boldsymbol{\theta}$. Also, recall that in all demand models considered in the paper, $\boldsymbol{A}_{ij}(\boldsymbol{\theta}) = 0$ for every $i \in Q(\boldsymbol{\theta})$ and $j \notin Q(\boldsymbol{\theta})$ (i.e. if a supplier has zero demand, then its price does not play a role in the demand equations of competitors). For a given $\boldsymbol{\theta}$ and a given $i \in Q(\boldsymbol{\theta})$, the constraints imposed by Eqs. (8) can be expressed as:

$$\sum_{j \in Q(\boldsymbol{\theta})} \boldsymbol{A}_{ij}(\boldsymbol{\theta}) p_j(\boldsymbol{\theta}) = \sum_{j \in Q(\boldsymbol{\theta})} \boldsymbol{A}_{ij}(\boldsymbol{\theta}) v_j(\boldsymbol{\theta}) \qquad (M_i(\boldsymbol{\theta}))$$

We refer to the constraint associated with the cost vector $\boldsymbol{\theta}$ and supplier $i \in Q(\boldsymbol{\theta})$ as $M_i(\boldsymbol{\theta})$. Any set of prices $\boldsymbol{p}(\boldsymbol{\theta})$ (for all $\boldsymbol{\theta} \in \Theta$) satisfying all these constraints implement the optimal allocations given by the solution of P_1.

In addition, by Corollary 4.1, we also need to guarantee that the expected interim transfers coincide with the optimal ones from P_1. We abuse notation and refer to the equality constraint on the expected transfers corresponding to supplier i and cost $\theta_i^j \in \Theta_i$ by $T_i(\theta_i^j)$. Recall that this constraint can be expressed as:

$$\sum_{\boldsymbol{\theta}_{-i} \in \Theta_{-i}} f_{-i}(\boldsymbol{\theta}_{-i}) x_i(\theta_i^j, \boldsymbol{\theta}_{-i}) p_i(\theta_i^j, \boldsymbol{\theta}_{-i}) \; = \; T_i(\theta_i^j) \qquad \forall i \in N, \; \forall \theta_i^j \in \Theta_i, \qquad (T_i(\theta_i^j))$$

Abusing notation, let $\boldsymbol{M}$ and $\boldsymbol{m}$ be the coefficient matrix and the corresponding RHS respectively defined by linear equations in ($M_i(\boldsymbol{\theta})$) and ($T_i(\theta_i^j)$), where each column is associated with a price $p_i(\boldsymbol{\theta})$ with $i \in Q(\theta)$.[41] The goal of the proof is to show that the system of linear equations given by $(\boldsymbol{M}, \boldsymbol{m})$ has a solution. Recall from Section 5.1 that the number of columns in $\boldsymbol{M}$ is greater than or equal to the number of rows. By the Rouché-Frobenius theorem, a system of linear equations $\boldsymbol{M}\boldsymbol{p} = \boldsymbol{m}$ is consistent (has a solution) if and only if the rank of its coefficient matrix $\boldsymbol{M}$ is equal to the rank of its augmented matrix $[\boldsymbol{M}|\boldsymbol{m}]$. To show whether the system of equations has a solution, we use an equivalent definition of consistency.

Lemma B.1 (Consistency of a system of linear equations)**.** *Consider the system of linear equations $\boldsymbol{M}\boldsymbol{p} = \boldsymbol{m}$. Let $\boldsymbol{M}_{i,*}$ denote the i^{th} row of $\boldsymbol{M}$. Then, the system is consistent (has a solution) if and only if for every vector $\boldsymbol{y}$ such that $\sum_i y_i \boldsymbol{M}_{i,*} = \boldsymbol{0}$, we have $\sum_i y_i m_i = 0$.*

To apply the above lemma, we define the associated coefficients as follows:

Definition B.1 (Associated Coefficients)**.** *For each row $M_i(\boldsymbol{\theta})$, let $a^i_{\boldsymbol{\theta}}$ denote the associated coefficient. Similarly, we denote by $b^i_{\theta_i^j}$ the coefficient associated to row $T_i(\theta_i^j)$. Let $(\boldsymbol{a}, \boldsymbol{b})$ be the vector of coefficients we just described.*

Rephrasing Lemma B.1 for our setting, for a system to be consistent we must have that for every vector $(\boldsymbol{a}, \boldsymbol{b})$ such that:

$$\sum_{\boldsymbol{\theta} \in \Theta} \sum_{\substack{i \in Q(\boldsymbol{\theta}) \\ i \neq \iota(Q(\boldsymbol{\theta}))}} a^i_{\boldsymbol{\theta}} M_i(\boldsymbol{\theta}) + \sum_{i \in N} \sum_{\theta_i^j \in \Theta_i} b^i_{\theta_i^j} T_i(\theta_i^j) \;=\; 0 \tag{12}$$

then the linear combination of the right hand side also equals zero, that is,

$$\sum_{\boldsymbol{\theta} \in \Theta} \sum_{\substack{i \in Q(\boldsymbol{\theta}) \\ i \neq \iota(Q(\boldsymbol{\theta}))}} a^i_{\boldsymbol{\theta}} \left(\sum_{j \in Q(\boldsymbol{\theta})} \boldsymbol{A}_{ij}(\boldsymbol{\theta}) v_j(\theta_j) \right) + \sum_{i \in N} \sum_{\theta_i^j \in \Theta_i} b^i_{\theta_i^j} \left(\theta_i^j X_i(\theta_i^j) + \sum_{k=j+1}^{|\Theta_i|} (\theta_i^k - \theta_i^{k-1}) X_i(\theta_i^k) \right) = 0. \tag{13}$$

Note that whenever the rows of $\boldsymbol{M}$ are linearly independent, the only vector of coefficients satisfying Eq. (12) is $(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{0}$ and, therefore, the system is trivially consistent.

B.1.1 Further Properties of the Coefficient Matrix

Through the rest of the section, we consider the general setting of Section 5.3. Given a matrix $\boldsymbol{A}$, we denote the i^{th} row of $\boldsymbol{A}$ by $\boldsymbol{A}_{i,*}$. Similarly, the j^{th} column is denoted by $\boldsymbol{A}_{*,j}$. For a subset of indices $Q \subset N$, $\boldsymbol{A}_Q$ denotes the principal submatrix of $\boldsymbol{A}$ obtained by selecting only the rows and

[41] Prices $p_i(\boldsymbol{\theta})$ with $i \notin Q(\boldsymbol{\theta})$ can safely be discarded, as all the coefficients of such columns are zero.

columns in Q. Similarly, $\boldsymbol{c}_Q$ denotes the vector obtained by selecting only the components in Q and $\mathbf{1}_Q$ denotes the vector of ones of dimension $|Q|$. We have the following result that characterizes an affine demand function for the set of active suppliers.

Lemma B.2. *Given a price vector $\boldsymbol{p}$ and the associated demand $\boldsymbol{d}(\boldsymbol{p})$, we denote by $Q = Q(\boldsymbol{p}) = \{i \in N : d_i(\boldsymbol{p}) > 0\}$. Then, demand $\boldsymbol{d}(\boldsymbol{p})$ can be expressed as:*

$$\boldsymbol{d}_Q(\boldsymbol{p}_Q) = (\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \left(\frac{1 - \mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{c}_Q - \boldsymbol{p}_Q\right)}{\mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\mathbf{1}_Q}\right)\mathbf{1}_Q\right). \tag{14}$$

Proof. We start by stating the KKT conditions for problem (LD($\boldsymbol{p}$)):

$$\begin{aligned} \boldsymbol{c} - \boldsymbol{D}\boldsymbol{x} - \boldsymbol{p} + \lambda\mathbf{1} + \boldsymbol{q} &= \mathbf{0} \\ \mathbf{1}'\boldsymbol{x} &= 1 \\ \boldsymbol{x} &\geq \mathbf{0} \\ \boldsymbol{x}'\boldsymbol{q} &= \mathbf{0} \\ \boldsymbol{q} &\geq 0, \end{aligned} \tag{15}$$

where λ is the multiplier associated to the equality constraint and $\boldsymbol{q}$ is the vector of multipliers associated to the non-negativity constraints. Define $\boldsymbol{v} = \boldsymbol{c} - \boldsymbol{D}\boldsymbol{x} - \boldsymbol{p} + \lambda\mathbf{1}$. By the KKT conditions we must have that $v_i = c_i - \boldsymbol{D}_{i,*}\boldsymbol{x} - p_i + \lambda = 0$, for all $i \in Q$. Therefore,

$$\mathbf{0} = \boldsymbol{v}_Q = \boldsymbol{c}_Q - \boldsymbol{D}_Q\boldsymbol{x}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q.$$

As $\boldsymbol{D}$ is positive definite and $\boldsymbol{D}_Q$ is a principal submatrix of $\boldsymbol{D}$ we have that $(\boldsymbol{D}_Q)^{-1}$ exists and, furthermore,

$$\boldsymbol{x}_Q = (\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q\right)$$

In addition, by the feasibility constraint, we must have $\mathbf{1}_Q'\boldsymbol{x}_Q = 1$ and hence,

$$1 = \mathbf{1}_Q'\boldsymbol{x}_Q = \mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q\right)$$

which implies

$$\lambda = \frac{1 - \mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{c}_Q - \boldsymbol{p}_Q\right)}{\mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\mathbf{1}_Q}\ .$$

Therefore,

$$\boldsymbol{x}_Q = (\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \left(\frac{1 - \mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{c}_Q - \boldsymbol{p}_Q\right)}{\mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\mathbf{1}_Q}\right)\mathbf{1}_Q\right) ,$$

as desired. □

The above demand specification exhibits a natural regularity property: if there is no demand for a particular product, the price of that product does not affect the demand for other products. In addition, it is simple to observe that any increase in price of a product with zero demand will not have an impact on the demand function either.

From Eq. (14), it should be clear that whenever two vector of prices $\boldsymbol{p}_Q$ and $\hat{\boldsymbol{p}}_Q$ satisfy

$$(\boldsymbol{D}_Q)^{-1}\left(\boldsymbol{p}_Q - \frac{\mathbf{1}_Q'(D_Q)^{-1}\boldsymbol{p}_Q}{\mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\mathbf{1}_Q}\mathbf{1}_Q\right) = (\boldsymbol{D}_Q)^{-1}\left(\hat{\boldsymbol{p}}_Q - \frac{\mathbf{1}_Q'(D_Q)^{-1}\hat{\boldsymbol{p}}_Q}{\mathbf{1}_Q'(\boldsymbol{D}_Q)^{-1}\mathbf{1}_Q}\mathbf{1}_Q\right), \tag{16}$$

we must have that $\boldsymbol{d}_Q(\boldsymbol{p}_Q) = \boldsymbol{d}_Q(\hat{\boldsymbol{p}}_Q)$. This observation is useful: it states that demands only depend on price differences. This freedom in setting unit prices is essential to our proof technique, as we will find unit prices that satisfy the same differences induced by the virtual costs and that simultaneously satisfy the expected interim transfer constraints.

By the previous observation, for $\boldsymbol{\theta} \in \Theta$ and each $i \in Q(\boldsymbol{\theta})$, the coefficient matrix $\boldsymbol{M}$ will consist of at most $Q(\boldsymbol{\theta})$ non-zero rows: $Q(\boldsymbol{\theta}) - 1$ correspond to the demand equations[42] and the remaining one corresponding to the expected transfer constraint. Note that for given $\boldsymbol{\theta} \in \Theta$, the demand equations are given by Eq. (16) where we replace Q by $Q(\boldsymbol{\theta})$ and $\boldsymbol{p}_Q$ by $\boldsymbol{p}_{Q(\boldsymbol{\theta})}(\boldsymbol{\theta})$ in the left hand side. In the right hand side we replace prices $\hat{\boldsymbol{p}}_Q$ by virtual costs $\boldsymbol{v}_{Q(\boldsymbol{\theta})}(\boldsymbol{\theta})$.

We now define the the demand submatrix associated to cost $\boldsymbol{\theta} \in \Theta$ as follows.

Definition B.2 (Demand submatrix of cost vector $\boldsymbol{\theta}$)**.** *For a given $\boldsymbol{\theta} \in \Theta$, we denote by $\boldsymbol{A}(\boldsymbol{\theta})$ the submatrix of $\boldsymbol{M}$ that contains the demand constraints for $\boldsymbol{\theta}$, that is, $\boldsymbol{A}(\boldsymbol{\theta})$ equals the left hand side of $(M_i(\boldsymbol{\theta}))_{i\in Q(\boldsymbol{\theta})}$.*

The following corollary of Lemma B.2 characterizes the matrix $\boldsymbol{A}(\boldsymbol{\theta})$ for the general affine demand models defined in Section 5.3. We include all demand equations in this matrix, even though as previously discussed, one of them is redundant.

[42]Note that if we can find prices p_Q satisfying the constraints imposed by $x_1, \ldots, x_{|Q|-1}$, then the last constraint will also be satisfied as $x_Q = 1 - \sum_{j=1}^{|Q|-1} x_j$.

Corollary B.1. *Let* $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{\theta}) = (\boldsymbol{D}_{Q(\boldsymbol{\theta})})^{-1}$*. Then, for every* $j \in Q(\boldsymbol{\theta})$ *and every* i *such that* $1 \leq i \leq Q(\boldsymbol{\theta})$*, the coefficient for* $p_j(\boldsymbol{\theta})$ *in equation* i *is given by:*

$$\boldsymbol{A}(\boldsymbol{\theta})_{ij} = -\boldsymbol{F}_{ij} + \frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})} \cdot \boldsymbol{F}_{*,j})(\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})})}{\mathbf{1}'_{Q(\boldsymbol{\theta})}\boldsymbol{F}\mathbf{1}_{Q(\boldsymbol{\theta})}}. \tag{17}$$

We now show that the associated demand vectors satisfy the original properties we wanted: the demand for a product is (weakly) decreasing in its own-price and (weakly) increasing in others' prices.

Lemma B.3 (Monotonicity)**.** *For every* $\boldsymbol{\theta} \in \Theta$ *and every* $i \in Q(\boldsymbol{\theta})$ *we have* $A(\boldsymbol{\theta})_{ii} < 0$ *and* $A(\boldsymbol{\theta})_{ij} \geq 0$ *for every* $j \in Q(\boldsymbol{\theta})$ *with* $j \neq i$*.*

Proof. We start by noting that, given the conditions imposed to matrix $\boldsymbol{\Gamma}$ in Section 5.3, we have that such matrix is a symmetric, non-singular, strictly-diagonally dominant M-matrix. In particular, an M-matrix with such properties has strictly positive diagonal elements, and non-positive off-diagonal elements. This proof uses several properties of M-matrices; the reader is referred to Horn and Johnson (1991) for the details.

Fix an arbitrary $\boldsymbol{\theta} \in \Theta$. By Corollary B.1, we have that

$$\boldsymbol{A}(\boldsymbol{\theta})_{ij} = -\boldsymbol{F}_{ij} + \frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})} \cdot \boldsymbol{F}_{*,j})(\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})})}{\mathbf{1}'_{Q(\boldsymbol{\theta})}\boldsymbol{F}\mathbf{1}_{Q(\boldsymbol{\theta})}},$$

where $\boldsymbol{F} = \boldsymbol{F}(\boldsymbol{\theta}) = (\boldsymbol{D}_{Q(\boldsymbol{\theta})})^{-1}$. The proof will consist of two steps. First, we argue that, if $\boldsymbol{F}$ is a symmetric, strictly diagonally dominant M-matrix, we have $A(\boldsymbol{\theta})_{ii} < 0$ and $A(\boldsymbol{\theta})_{ij} \geq 0$ for every $i, j \in Q(\boldsymbol{\theta})$ with $j \neq i$ as desired. Second, we show that $\boldsymbol{F}$ is indeed a symmetric, strictly diagonally dominant M-matrix.

To that end, suppose $\boldsymbol{F}$ is a symmetric, strictly diagonally dominant M-matrix. Then, $\boldsymbol{F}_{ii} > 0$ and $\boldsymbol{F}_{ij} \leq 0$ and we must have that, for every row, the sum of the elements in a row must be strictly positive. By symmetry, this is true also for the sum of the elements in a column. In turn, this implies that the sum of all elements in the matrix is strictly positive and hence,

$$\boldsymbol{A}(\boldsymbol{\theta})_{ij} = -\boldsymbol{F}_{ij} + \frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})} \cdot \boldsymbol{F}_{*,j})(\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})})}{\mathbf{1}'_{Q(\boldsymbol{\theta})}\boldsymbol{F}\mathbf{1}_{Q(\boldsymbol{\theta})}} < -\boldsymbol{F}_{ij} + \boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})} \tag{18}$$

Then, we have $\boldsymbol{A}(\boldsymbol{\theta})_{ii} < -\boldsymbol{F}_{ii} + \boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})} \leq 0$, where the last inequality follows because $F_{ij} \leq 0$, $\forall i \neq j$. Hence $\boldsymbol{A}(\boldsymbol{\theta})_{ii} < 0$ as desired. Similarly, $\boldsymbol{A}(\boldsymbol{\theta})_{ij} \geq 0$ follows from the fact that both terms in the summation are non-negative. Therefore, we have shown that if $\boldsymbol{F}$ is a symmetric, strictly diagonally dominant M-matrix, the result follows.

To complete the proof, we show that $\boldsymbol{F}$ satisfies the stated properties. As usual, for the given $\boldsymbol{\theta} \in \Theta$, let $Q = Q(\boldsymbol{\theta}) = \{k \in N : x_k(\boldsymbol{\theta}) > 0\}$ and $\overline{Q} = \overline{Q}(\boldsymbol{\theta}) = \{k \in N : x_k(\boldsymbol{\theta}) = 0\}$. From now on, we omit the dependence on $\boldsymbol{\theta}$ to simplify notation. By the KKT conditions (Eq. (15)), we have that $\boldsymbol{x} = \boldsymbol{D}^{-1}(\boldsymbol{c} - \boldsymbol{p} + \lambda \mathbf{1} + \boldsymbol{q})$. (Recall that $\boldsymbol{D}^{-1} = \boldsymbol{\Gamma}$.) By definition, we have that

$$\mathbf{0} = \boldsymbol{x}_{\overline{Q}} = \boldsymbol{\Gamma}_{\overline{Q},*}(\boldsymbol{c} - \boldsymbol{p} + \lambda\mathbf{1} + \boldsymbol{q}) = \boldsymbol{\Gamma}_{\overline{Q},\overline{Q}}(\boldsymbol{c}_{\overline{Q}} - \boldsymbol{p}_{\overline{Q}} + \lambda\mathbf{1}_{\overline{Q}} + \boldsymbol{q}_{\overline{Q}}) + \boldsymbol{\Gamma}_{\overline{Q},Q}(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q),$$

where we used $\boldsymbol{q}_Q = 0$ by the KKT conditions. Note that $\boldsymbol{\Gamma}_{\overline{Q},\overline{Q}}$ is a principal submatrix of a non-singular M-matrix. Thus, $(\boldsymbol{\Gamma}_{\overline{Q},\overline{Q}})^{-1}$ exists and:

$$(\boldsymbol{c}_{\overline{Q}} - \boldsymbol{p}_{\overline{Q}} + \lambda\mathbf{1}_{\overline{Q}} + \boldsymbol{q}_{\overline{Q}}) = -(\boldsymbol{\Gamma}_{\overline{Q},\overline{Q}})^{-1}\boldsymbol{\Gamma}_{\overline{Q},Q}(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q) \tag{19}$$

In addition, we have

$$\begin{aligned}
\boldsymbol{x}_Q &= \boldsymbol{\Gamma}_{Q,*}(\boldsymbol{c} - \boldsymbol{p} + \lambda\mathbf{1} + \boldsymbol{q}) \\
&= \boldsymbol{\Gamma}_{Q,Q}(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q) + \boldsymbol{\Gamma}_{Q,\overline{Q}}(\boldsymbol{c}_{\overline{Q}} - \boldsymbol{p}_{\overline{Q}} + \lambda\mathbf{1}_{\overline{Q}} + \boldsymbol{q}_{\overline{Q}}) \\
&= \boldsymbol{\Gamma}_{Q,Q}(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q) - \boldsymbol{\Gamma}_{Q,\overline{Q}}(\boldsymbol{\Gamma}_{\overline{Q},\overline{Q}})^{-1}\boldsymbol{\Gamma}_{\overline{Q},Q}(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q) \\
&= \left(\boldsymbol{\Gamma}_{Q,Q} - \boldsymbol{\Gamma}_{Q,\overline{Q}}(\boldsymbol{\Gamma}_{\overline{Q},\overline{Q}})^{-1}\boldsymbol{\Gamma}_{\overline{Q},Q}\right)(\boldsymbol{c}_Q - \boldsymbol{p}_Q + \lambda\mathbf{1}_Q),
\end{aligned}$$

where the second to last equality follows from Eq. (19). By the definition of $\boldsymbol{D}_Q$, we have $(\boldsymbol{D}_Q)^{-1} = \left(\boldsymbol{\Gamma}_{Q,Q} - \boldsymbol{\Gamma}_{Q,\overline{Q}}(\boldsymbol{\Gamma}_{\overline{Q},\overline{Q}})^{-1}\boldsymbol{\Gamma}_{\overline{Q},Q}\right)$. In turn, this implies that $(\boldsymbol{D}_Q)^{-1}$ is the Schur complement of $\boldsymbol{\Gamma}_{\overline{Q},\overline{Q}}$ in $\boldsymbol{\Gamma}$. In particular, we have that the Schur complement of a M-matrix is also a M-matrix, and non-singularity, symmetry and strict diagonal dominance are preserved in Schur complementation (Carlson and Markham 1979, Horn and Johnson 1991). Therefore, the matrix $\boldsymbol{F}$ satisfies the desired properties, which completes the proof. $\square$

Next, we establish two useful properties on the demand submatrices associated to cost profiles $\boldsymbol{\theta}$ such that all agents are active in $\boldsymbol{\theta}$.

Lemma B.4. *Let $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{\theta})$ for any $\boldsymbol{\theta} \in \Theta$ such that $Q(\boldsymbol{\theta}) = N$ be as defined by Corollary B.1. Then, $\boldsymbol{A}$ is symmetric and has rank $n-1$.*

Proof. Note that, whenever $Q(\boldsymbol{\theta}) = N$, we have $\boldsymbol{F} = \boldsymbol{D}^{-1}$ where $\boldsymbol{F}$ is as defined in Claim B.1. By assumption, $\boldsymbol{D}^{-1}$ is symmetric and positive definite. Therefore, $\boldsymbol{A}$ is also symmetric by definition.

To show that $\boldsymbol{A}$ has rank $n-1$, let $\boldsymbol{I}$ denote the identity matrix of size n. Note that $\boldsymbol{A} =$

$\boldsymbol{D}^{-1}\left(-\boldsymbol{I}+\boldsymbol{1}\frac{\boldsymbol{1}'\boldsymbol{D}^{-1}}{\boldsymbol{1}'\boldsymbol{D}^{-1}\boldsymbol{1}}\right)$. Therefore,

$$rank(\boldsymbol{A}) \geq rank(\boldsymbol{D}^{-1}) + rank\left(-\boldsymbol{I}+\boldsymbol{1}\frac{\boldsymbol{1}'\boldsymbol{D}^{-1}}{\boldsymbol{1}'\boldsymbol{D}^{-1}\boldsymbol{1}}\right) - n = rank\left(-\boldsymbol{I}+\boldsymbol{1}\frac{\boldsymbol{1}'\boldsymbol{D}^{-1}}{\boldsymbol{1}'\boldsymbol{D}^{-1}\boldsymbol{1}}\right),$$

as $\boldsymbol{D}^{-1}$ has full rank. In addition, we have[43]

$$rank\left(-\boldsymbol{I}+\boldsymbol{1}\frac{\boldsymbol{1}'\boldsymbol{D}^{-1}}{\boldsymbol{1}'\boldsymbol{D}^{-1}\boldsymbol{1}}\right) \geq \left|n - rank\left(\boldsymbol{1}\frac{\boldsymbol{1}'\boldsymbol{D}^{-1}}{\boldsymbol{1}'\boldsymbol{D}^{-1}\boldsymbol{1}}\right)\right| \geq n-1,$$

as the matrix $\boldsymbol{1}\frac{\boldsymbol{1}'\boldsymbol{D}^{-1}}{\boldsymbol{1}'\boldsymbol{D}^{-1}\boldsymbol{1}}$ has rank exactly one. The converse follows just from the definition of $\boldsymbol{A}$, as we know that one row must be redundant as all demands must some up to one. □

In order to show that the system of equations is consistent, we want to find prices $\boldsymbol{p}$ such that $\boldsymbol{x}(\boldsymbol{p}) = \boldsymbol{x}(v(\boldsymbol{\theta}))$, where $v(\boldsymbol{\theta}) = (v_1(\boldsymbol{\theta}), \ldots, v_n(\boldsymbol{\theta}))$ is defined as the vector of virtual costs. That is, we must have $\boldsymbol{A}\boldsymbol{p} = \boldsymbol{A}v(\boldsymbol{\theta})$. Therefore, Lemma B.4 states that for $\boldsymbol{\theta} \in \Theta$ such that $Q(\boldsymbol{\theta}) = N$, the dimension of prices satisfying those demand constraints is exactly one, as $\boldsymbol{A}$ has rank $n-1$.

B.1.2 Coefficient Matrix for Hotelling Model

We provide a brief note on the Hotelling model. While all the material in this section is presented with the general affine demand model in mind to avoid cumbersome notation, we now show that all the properties of matrix $\boldsymbol{A}(\boldsymbol{\theta})$ shown above (that we use to prove our main result) also hold for the Hotelling Model.

Remark B.1. *Given $\boldsymbol{\theta}$, let $Q(\boldsymbol{\theta})$ be the set of active agents ordered from leftmost to rightmost. Then,*

$$\boldsymbol{A}(\boldsymbol{\theta}) = \frac{1}{2\delta}\begin{bmatrix} -1 & 1 & 0 & 0 & \ldots & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \ldots & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 1 & -1 \end{bmatrix}$$

This follows from the fact that, in the Hotelling model, suppliers split the market with their immediate active neighbors; in particular, i obtains $\frac{p_j - p_i + \delta|\ell_j - \ell_i|}{2\delta}$ units from the segment $[\ell_i, \ell_j]$ and j the rest. If i is the leftmost active supplier, he obtains all the demand in the $[0, \ell_i]$ segment; similarly, if he is the rightmost active supplier, he obtains all the demand in $[\ell_i, 1]$. Renaming the suppliers in $Q(\boldsymbol{\theta})$ as $1, 2, \ldots$, with numbers increasing from left to right suppliers, we have that the

[43]Matrix property: $rank(A - B) \geq |rank(A) - rank(B)|$

demand for the leftmost active supplier 1 is $\ell_1 + \frac{p_2 - p_1 + \delta|\ell_2 - \ell_1|}{2\delta}$. Note that the coefficient of the prices in this equation are represented by the first row of the matrix. Similarly, the demand for supplier 2 is $\frac{p_1 - p_2 + \delta|\ell_2 - \ell_1|}{2\delta} + \frac{p_3 - p_2 + \delta|\ell_3 - \ell_2|}{2\delta}$; this is summarized by the second row, and so on.

It is immediate to see that, under the Hotelling model, Lemmas B.3 and B.4 hold. Similarly, for a fixed set Q, the demands only depend on price differences and not on actual prices.

B.2 Definitions and Notation

We now state some definitions that we will use to prove the main theorem. Let $\underline{\theta_i}$ and $\overline{\theta_i}$ denote the lowest and highest values in Θ_i. For each $j \in N$, let θ_j^u be the maximum $\theta_j \in \Theta_j$ under which there exists a profile $\boldsymbol{\theta} = (\theta_j, \boldsymbol{\theta}_{-j})$ such that $j \in Q(\boldsymbol{\theta})$. We may assume that $\underline{\theta_j} \leq \theta_j^u$ for all agents $j \in N$, as otherwise we can consider (w.l.o.g.) the reduced problem in which all agents for which the condition is violated are removed. In addition, note that for agent j all constraints and coefficients associated to $\theta_j > \theta_j^u$ will not play a role in our analysis, because agent j is inactive for all profiles with $\theta_j > \theta_j^u$.

Two profiles $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \Theta$ are defined to be *adjacent* if and only if $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ only differ in one component and $Q(\boldsymbol{\theta}) = Q(\boldsymbol{\theta}')$, where $Q(\boldsymbol{\theta})$ is the set of active firms in the relaxed optimal solution under profile $\boldsymbol{\theta}$. Given two profiles $\boldsymbol{\theta}, \boldsymbol{\theta}'$, we define $\boldsymbol{\theta}$ to be *reachable* from $\boldsymbol{\theta}'$ if there exists a sequence of profile $\{\boldsymbol{\theta}_0 = \boldsymbol{\theta}', \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_K = \boldsymbol{\theta}\}$ such that $Q(\boldsymbol{\theta}_k) \subseteq Q(\boldsymbol{\theta}_{k-1})$ for all $1 \leq k \leq K$, and the sub-profiles $(\boldsymbol{\theta}_{k-1})_{Q(\boldsymbol{\theta}_k)}$ and $(\boldsymbol{\theta}_k)_{Q(\boldsymbol{\theta}_k)}$ differ in at most one component; that is, at most one agent among those active in $\boldsymbol{\theta}_k$ has a different cost in both profiles.

Definition B.3 (Acceptable set)**.** *A subset of profiles $\tilde{\Theta} \subseteq \Theta$ is an* acceptable set *if the following conditions are simultaneously satisfied:*

1. *$Q(\boldsymbol{\theta}) = N$ for every $\boldsymbol{\theta} \in \tilde{\Theta}$.*

2. *For each agent i, let $\tilde{\Theta}_i = \{\theta_i \in \Theta_i : \exists \boldsymbol{\theta}_{-i}$ such that $(\theta_i, \boldsymbol{\theta}_{-i}) \in \tilde{\Theta}\}$. Then, for every $\theta_i \in \Theta_i$ such that $\min \tilde{\Theta}_i \leq \theta_i \leq \max \tilde{\Theta}_i$ we must have $\theta_i \in \tilde{\Theta}_i$. That is, each $\tilde{\Theta}_i$ must be a (discrete) interval.*

3. *For every profile $\boldsymbol{\theta}$ such that $\theta_i \in \tilde{\Theta}_i$ for all $i \in N$, we must have $\boldsymbol{\theta} \in \tilde{\Theta}$.*

We abuse notation to denote $\min \tilde{\Theta}_i = \min\{\theta_i : \theta_i \in \tilde{\Theta}_i\}$ and $\max \tilde{\Theta}_i = \max\{\theta_i : \theta_i \in \tilde{\Theta}_i\}$. The above definition of acceptable set will help us characterize sufficient conditions under which the optima of the relaxed and original problems agree. In particular, let a market be defined by the set of suppliers, their product characteristics and cost distributions, as well as the demand model. We define a *relaxation-is-optimal market* (RIOM) as follows.

Definition B.4 (RIOM)**.** *A market is* relaxation-is-optimal market *(RIOM) if there exists an acceptable set* $\tilde{\Theta}$ *under which the following (additional) conditions are satisfied:*

4. *For every* $i \in N$ *we have* $|\tilde{\Theta}_i| \geq 3$*.*

5. *For all* $i \in N$ *and* θ_i *such that* $\max \tilde{\Theta}_i \leq \theta_i \leq \theta_i^u$*, there exists a profile* $\boldsymbol{\theta} = (\theta_i, \theta_{-i})$ *with* $i \in Q(\boldsymbol{\theta})$ *and a profile* $\boldsymbol{\theta}' \in \tilde{\Theta}$ *such that profile* $\boldsymbol{\theta}$ *is reachable from* $\boldsymbol{\theta}'$*.*

Intuitively, a market will be RIOM if (1) there exists a solution in which all agents are active, and (2) the difference in virtual costs between adjacent points in the support is "small enough". If the difference between adjacent virtual costs is small, then by changing a cost by the following (or preceding) one, we do not expect the allocation (and hence the set of active suppliers) to change much. Therefore, Conditions (2) and (3) will be satisfied. Similarly, if there exists a cost profile $\boldsymbol{\theta}$ for which all agents are active, one would expect that this will also be true for the cost profiles close to $\boldsymbol{\theta}$ provided adjacent virtual costs are close enough. Therefore, Condition (4) will be satisfied. Finally, a small difference between adjacent virtual costs also implies Condition (5); we can change one cost at a time by an adjacent one while having some control over the set of active suppliers, and therefore we can construct a path of profiles that can take us from $\boldsymbol{\theta}'$ to $\boldsymbol{\theta}$.

Our main theorem will state that, if the market is RIOM, then we have that the optima of the original and relaxed problems agree and thus we can characterize the optimal mechanism. Therefore, we now show that the conditions of Theorem 5.1 for the Hotelling model and Theorem 5.2 for the general affine model imply that the markets are RIOM.

Lemma B.5. *Any market satisfying the conditions of Theorem 5.1 is RIOM.*

Proof. Recall that all firms are active in the relaxed optimal solution under profile θ, if $|v_j(\theta_j) - v_i(\theta_i)| \leq \delta|\ell_j - \ell_i|$, for all $i, j \in N$. Hence, by Condition (1) in the statement of the theorem, a profile $\boldsymbol{\theta}$ in which $Q(\boldsymbol{\theta}) = N$ must exist. Furthermore, $|v_{i+1}(\theta_i) - v_i(\theta_i)| \leq \delta(\ell_{i+1} - \ell_i)/4$ for all $i \in N$. Condition (2) in the statement of the theorem states that $v_i(\theta_i^{j+1}) - v_i(\theta_i^j) \leq \frac{\delta c^*}{8}$ for all $i \in N$, and $\theta_i^j \in \Theta_i$. Using the two conditions it is simple to show that, by letting θ_i^k denote θ_i, we must have $Q(\theta_i^{k+2}, \boldsymbol{\theta}_{-i}) = Q(\theta_i^{k-2}, \boldsymbol{\theta}_{-i}) = N$, provided these exist. Further, let $\boldsymbol{\theta} = (\theta_i, \boldsymbol{\theta}_{-i})$ and define $\tilde{\Theta}_i = \{\theta_i' \in \Theta_i : |v_i(\theta_i') - v_i(\theta_i)| \leq \frac{\delta c^*}{4}\}$. As $|\Theta_i| \geq 3$, we must have that $|\tilde{\Theta}_i| \geq 3$. In addition, notice that for every pair $\theta_i', \theta_j' \in \tilde{\Theta}_i \times \tilde{\Theta}_j$ we have that $|v_i(\theta_i') - v_j(\theta_j')| \leq |v_i(\theta_i) - v_j(\theta_j)| + \frac{\delta c^*}{2} \leq \frac{3|\ell_i - \ell_j|}{4}\delta$. Therefore, defining $\tilde{\Theta} = \prod_{i \in N} \tilde{\Theta}_i$, we get that $Q(\theta) = N$, $\forall \theta \in \tilde{\theta}$, and $\tilde{\Theta}$ is an acceptable set satisfying Condition (4). Finally, we show that the reachability requirement (Condition (5)) is satisfied.

To that end, we explicitly construct a sequence of profiles that are reachable from a $\boldsymbol{\theta}' \in \tilde{\Theta}$, and such that for all $i \in N$ and all θ_i with $\max \tilde{\Theta}_i \leq \theta_i \leq \theta_i^u$, there exists a profile in the sequence of profiles for which i has cost θ_i and is active. Let $\boldsymbol{\theta}_0$ be a profile such that $(\boldsymbol{\theta}_0)_i = \max \tilde{\Theta}_i$ for all $i \in N$. Note that $\boldsymbol{\theta}_0 \in \tilde{\Theta}$ by construction. From $\boldsymbol{\theta}_0$, we construct a profile $\boldsymbol{\theta}_1$ by selecting the agent j with the minimum virtual cost, and increasing his cost to the adjacent one, call it $(\boldsymbol{\theta}_1)_j$. The costs of all other agents do not change in the new profile. Note that, for all $i \neq j$:

$$v_j((\boldsymbol{\theta}_1)_j) - v_i((\boldsymbol{\theta}_1)_i) = v_j((\boldsymbol{\theta}_1)_j) - v_j((\boldsymbol{\theta}_0)_j) + v_j((\boldsymbol{\theta}_0)_j) - v_i((\boldsymbol{\theta}_1)_i) \leq \delta c^*/8 \, , \tag{20}$$

because the difference between adjacent virtual costs is bounded by $\delta c^*/8$ by assumption, and $v_j((\boldsymbol{\theta}_0)_j) \leq v_i((\boldsymbol{\theta}_1)_i) = v_i((\boldsymbol{\theta}_0)_i)$, by construction. Hence, agent j remains active and all other agents remain active by monotonicity of Hotelling demand. We can inductively apply this procedure —select the agent with lowest virtual cost and increase his cost to the adjacent one— to obtain profiles that are adjacent and in which all agents are active.

Eventually, we will reach a profile $\boldsymbol{\theta}_K$ for which we cannot increase the cost of the agent j with lowest virtual cost; this means that $(\boldsymbol{\theta}_K)_j = \overline{\theta_j}$. Further, it must be that $\theta^u_j = \overline{\theta_j}$. Thus, we have shown that for all θ_j with $\max \tilde{\Theta}_j \leq \theta_j \leq \theta^u_j$, there exists a profile in the sequence of profiles for which j has cost θ_j, j is active, and such profile is reachable from $\boldsymbol{\theta}_0$.

Let $U = \{j\}$; from now on, the set U will contain all agents who have reached θ^u. Construct a profile $\boldsymbol{\theta}_{K+1}$ by selecting the agent j' with the lowest virtual cost among those in $N \backslash U$, and increasing his virtual cost to the adjacent one. Now three possibilities arise:

1. If the cost of agent j' can be increased and j' remains active, then we just increase his cost and repeat.

2. If the cost of such agent cannot be increased further, this implies that we have shown our claim for j', because we have reached $\overline{\theta_{j'}}$; hence, we can add him to U and repeat.

3. Finally, we consider the case in which the cost of j' can be increased but in doing so we have $j' \notin Q(\boldsymbol{\theta}_{K+1})$; then, we must have $\theta^u_{j'} = (\boldsymbol{\theta}_K)_{j'}$. To see why this holds, note that as j' is the agent with lowest virtual cost among the ones in $N \backslash U$, he can only be inactive in the new profile if an agent in U (agent j) grabs the demand j' had in the old profile $(\boldsymbol{\theta}_K)$ (by a similar argument to equation (20)). As a consequence, it is simple to observe that agent j will keep j' inactive even if the virtual costs of other agents increase. This together with $(\boldsymbol{\theta}_K)_j = \overline{\theta}_j$ shows our claim for j'.

We proceed by adding j' to U and defining $\boldsymbol{\theta}_{K+1} = (\overline{\theta_{j'}}, (\boldsymbol{\theta}_K)_{-j'})$; by construction, $\boldsymbol{\theta}_{K+1}$ is reachable from $\boldsymbol{\theta}_0$. We conclude the proof by noting that we can inductively apply this procedure. Each time we add an agent to U, we have shown the claim for such agent. Specifically, every time the cost of an agent cannot be increased because he will become inactive, it must be caused by the fact that one of the costs of at least one agent in U is preventing for doing so. However, in the current profile all agents in U are at their maximum costs by construction; thus, such agent has reached θ^u and we have shown that the statement is true for him as well. $\square$

We now show an analogous result for the case of affine demand models. Let $\overline{\boldsymbol{\theta}} = (\overline{\theta_j})_{j\in N}$. To prove the following lemma we will assume that $Q(\overline{\boldsymbol{\theta}}) = N$, that is all suppliers are active in the relaxed optimal solution at the largest cost profile. This assumption significantly simplifies the proof and the notation required. However, one can also show that any market satisfying the conditions of Theorem 5.2 is RIOM even if $Q(\overline{\boldsymbol{\theta}}) \subset N$.

Lemma B.6. *Any market satisfying the conditions of Theorem 5.2 and for which $Q(\overline{\boldsymbol{\theta}}) = N$ is RIOM.*

Proof. First, note that the existence of d^* and Condition (1) in the statement of the theorem defines an acceptable set $\tilde{\Theta}$. Furthermore, note that Condition (2) in the statement of the theorem implies that Condition (4) in the definition of RIOM will be satisfied. Finally, Condition (5) in the definition of RIOM is trivially satisfied because $\overline{\boldsymbol{\theta}} \in \tilde{\Theta}$. $\square$

Furthermore, in the setting of Lemma B.6 we can also characterize d^* as follows. Let $M = \max_{j\in N} |\tilde{\Theta}_j|$, and let $\boldsymbol{A} = \boldsymbol{A}(\overline{\boldsymbol{\theta}})$ be the coefficient matrix associated with profile $\overline{\boldsymbol{\theta}}$, as defined in Definition B.2. Then, as long as $d^* < \frac{2}{M} \min_{i\in N}\{(-\frac{1}{\boldsymbol{A}_{ii}} x_i(\overline{\boldsymbol{\theta}}))\}$, we have that the market is RIOM. Note that $\boldsymbol{x}(\overline{\boldsymbol{\theta}})$, the optimal allocations for the relaxed problem at profile $\overline{\boldsymbol{\theta}}$, depend on the models primitives through the demand system and the virtual costs.

B.3 Auxiliary Lemma

We state and prove the following Lemma, which will play a key role in the proof of the main theorem.

Lemma B.7. *Suppose the coefficients $(\boldsymbol{a}, \boldsymbol{b})$ are such that equality in Eq. (12) holds. For each $i \in N$ and each $\theta_i \in \Theta_i$, let $g_i(\theta_i)$ be defined as $g_i(\theta_i) = \frac{b^i_{\theta_i}}{f_i(\theta_i)}$. Then for each $\boldsymbol{\theta} \in \Theta$, we must have*

$$\sum_{i\in Q(\boldsymbol{\theta})} g_i(\theta_i) x_i(\boldsymbol{\theta}) = 0 \tag{21}$$

Proof. Fix $\boldsymbol{\theta} \in \Theta$. We show the result for the general affine demand model as described in Section 5.3. Recall that the coefficients of the matrix corresponding to the demand equations (that is, Eqs. $(M_i(\boldsymbol{\theta}))$) are as defined by Eq. (17). As the equality in Eq. (12) holds, for each $j \in Q(\boldsymbol{\theta})$ we must have:

$$b^j_{\theta_j} f(\boldsymbol{\theta}_{-j}) x_j(\boldsymbol{\theta}) + \sum_{i=1}^{Q(\boldsymbol{\theta})-1} a^i_{\boldsymbol{\theta}} \left(-\boldsymbol{F}_{ij} + \frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})} \cdot \boldsymbol{F}_{*,j})(\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})})}{\mathbf{1}'_{Q(\boldsymbol{\theta})} \boldsymbol{F} \mathbf{1}_{Q(\boldsymbol{\theta})}} \right) = 0,$$

where we have used the fact that one constraint is indeed redundant (and thus the summation goes to $Q(\boldsymbol{\theta}) - 1$ instead of $Q(\boldsymbol{\theta})$). Therefore,

$$\begin{aligned}
\sum_{j \in Q(\boldsymbol{\theta})} b^j_{\theta_j} f(\boldsymbol{\theta}_{-j}) x_j(\boldsymbol{\theta}) &= -\sum_{j \in Q(\boldsymbol{\theta})} \sum_{i=1}^{Q(\boldsymbol{\theta})-1} a^i_{\boldsymbol{\theta}} \left(-\boldsymbol{F}_{ij} + \frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})} \cdot \boldsymbol{F}_{*,j})(\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})})}{\mathbf{1}'_{Q(\boldsymbol{\theta})} \boldsymbol{F} \mathbf{1}_{Q(\boldsymbol{\theta})}} \right) \\
&= -\sum_{i=1}^{Q(\boldsymbol{\theta})-1} a^i_{\boldsymbol{\theta}} \left(\sum_{j \in Q(\boldsymbol{\theta})} \left(-\boldsymbol{F}_{ij} + \frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})} \cdot \boldsymbol{F}_{*,j})(\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})})}{\mathbf{1}'_{Q(\boldsymbol{\theta})} \boldsymbol{F} \mathbf{1}_{Q(\boldsymbol{\theta})}} \right) \right) \\
&= -\sum_{i=1}^{Q(\boldsymbol{\theta})-1} a^i_{\boldsymbol{\theta}} \left(-\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})} + \boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})} \left(\sum_{j \in Q(\boldsymbol{\theta})} \frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})} \cdot \boldsymbol{F}_{*,j})}{\mathbf{1}'_{Q(\boldsymbol{\theta})} \boldsymbol{F} \mathbf{1}_{Q(\boldsymbol{\theta})}} \right) \right) \\
&= -\sum_{i=1}^{Q(\boldsymbol{\theta})-1} a^i_{\boldsymbol{\theta}} \left(-\boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})} + \boldsymbol{F}_{i,*} \cdot \mathbf{1}_{Q(\boldsymbol{\theta})} \right) \\
&= 0
\end{aligned}$$

To complete the proof, note that $\sum_{j \in Q(\boldsymbol{\theta})} b^j_{\theta_j} f(\boldsymbol{\theta}_{-j}) x_j(\boldsymbol{\theta}) = f(\boldsymbol{\theta}) \left(\sum_{j \in Q(\boldsymbol{\theta})} g_j(\theta_j) x_j(\boldsymbol{\theta}) \right) = 0$. Hence, $\sum_{j \in Q(\boldsymbol{\theta})} g_j(\theta_j) x_j(\boldsymbol{\theta}) = 0$ as desired. □

B.4 Main Theorem

We can now state and prove our main theorem.

Theorem B.1. *Consider the general affine demand model in which agents have arbitrary costs distributions. If the market is RIOM, then $OPT(P_0) = OPT(P_1)$.*

Proof. To show $OPT(P_0) = OPT(P_1)$, we show that the system of equations is consistent. Let $(\boldsymbol{a}, \boldsymbol{b})$ be a vector of coefficients satisfying Eq. (12). Let $g_i(\theta_i)$ be as defined in the statement of Lemma B.7. The idea of the proof is to first show that, if a market is RIOM, then all $g_i(\theta_i)$ must be zero. Then, we show that if $g_i(\theta_i) = 0$, for all $\theta_i \in \Theta_i$ and all $i \in N$, then the system is consistent and thus $OPT(P_0) = OPT(P_1)$ as desired. Consequently, the proof is divided into the following steps:

Step 1: Show that if $(\boldsymbol{a}, \boldsymbol{b})$ satisfies Eq. (12) and a market is RIOM all $g_i(\theta_i)$ must be zero. Let $\tilde{\Theta} \subseteq \Theta$ be such that it satisfies Conditions (1)-(5) in Definitions B.3 and B.4 respectively (we know such $\tilde{\Theta}$ exists as the market is RIOM) . Step 1 is further divided into the following two sub-steps:

(a) **Step 1.a:** Show that $g_i(\theta_i) = 0$ for all $\theta_i \in \tilde{\Theta}_i$ and all $i \in N$.

(b) **Step 1.b:** Show that $g_i(\theta_i) = 0$ for all $\theta_i \notin \tilde{\Theta}_i$ and all $i \in N$.

Step 2: Show that $g_i(\theta_i) = 0$, for all $\theta_i \in \Theta_i$ and all $i \in N$, implies consistency of the system of linear equations.

Step 1.a : Show $g_i(\theta_i) = 0$, for all $\theta_i \in \tilde{\Theta}_i$ and all $i \in N$. By assumption, $\tilde{\Theta}$ satisfies conditions (1)-(5). Therefore, for every $\boldsymbol{\theta} \in \tilde{\Theta}$ we must have $Q(\boldsymbol{\theta}) = N$ (by Condition (1)). Consider two profiles $\boldsymbol{\theta} = (\theta_i, \boldsymbol{\theta}_{-i})$ and $\boldsymbol{\theta}' = (\theta_i', \boldsymbol{\theta}_{-i})$ which only differ in agent i's cost and such that $\boldsymbol{\theta}, \boldsymbol{\theta}' \in \tilde{\Theta}$. By the definition of $\tilde{\Theta}$, such pair of profiles exists (Conditions (3) and (4)). By Eq. (21), we must have $g_i(\theta_i)x_i(\boldsymbol{\theta}) + \sum_{j \neq i} g_j(\theta_j)x_j(\boldsymbol{\theta}) = 0$ and $g_i(\theta_i')x_i(\boldsymbol{\theta}') + \sum_{j \neq i} g_j(\theta_j)x_j(\boldsymbol{\theta}') = 0$. Hence, by subtracting the second equality from the first one we obtain

$$g_i(\theta_i)x_i(\boldsymbol{\theta}) - g_i(\theta_i')x_i(\boldsymbol{\theta}') = \sum_{j \neq i} g_j(\theta_j) \left[x_j(\boldsymbol{\theta}') - x_j(\boldsymbol{\theta})\right].$$

For each $j \in N$, we must have $x_j(\boldsymbol{\theta}') - x_j(\boldsymbol{\theta}) = \boldsymbol{A}(\boldsymbol{\theta})_{ji} \left(v_i(\theta_i') - v_i(\theta_i)\right)$, where we used the fact that $\boldsymbol{A}(\boldsymbol{\theta}) = \boldsymbol{A}(\boldsymbol{\theta}')$ by definition, as the same set of agents are active. Let $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{\theta})$, and note that this $\boldsymbol{A}$ agrees with the one in Lemma B.4, because $Q(\boldsymbol{\theta}) = N$. Hence, we can re-write the above equality as:

$$g_i(\theta_i)x_i(\boldsymbol{\theta}) - g_i(\theta_i')x_i(\boldsymbol{\theta}') = \big(v_i(\theta_i') - v_i(\theta_i)\big) \left(\sum_{j \neq i} g_j(\theta_j)\boldsymbol{A}_{ji}\right),$$

and therefore,

$$\frac{g_i(\theta_i)x_i(\boldsymbol{\theta}) - g_i(\theta_i')x_i(\boldsymbol{\theta}')}{v_i(\theta_i') - v_i(\theta_i)} = \left(\sum_{j \neq i} g_j(\theta_j)\boldsymbol{A}_{ji}\right). \tag{22}$$

Fix an arbitrary $j \in N$ with $j \neq i$ and $\boldsymbol{A}_{ij} \neq 0$. By strict diagonal dominance of $\boldsymbol{F}$, such j always exists (see Eq. (18)). Let θ_j be the cost of agent j in both $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$, where the cost profiles are as defined above. Let $\theta_j' \in \Theta_j$ be such that $\theta_j' \neq \theta_j$ and $\theta_j' \in \tilde{\Theta}_j$ (by Conditions (3) and (4), such θ_j' exists). Define $\tilde{\boldsymbol{\theta}} = (\theta_i, \theta_j', \boldsymbol{\theta}_{-i,j})$ and $\tilde{\boldsymbol{\theta}}' = (\theta_i', \theta_j', \boldsymbol{\theta}_{-i,j})$. The only thing we assumed about θ_j was $\theta_j \in \tilde{\Theta}_j$. Therefore, the above equality must also hold for any $\tilde{\theta}_j \in \tilde{\Theta}_j$. That is,

$$\frac{g_i(\theta_i)x_i(\tilde{\boldsymbol{\theta}}) - g_i(\theta_i')x_i(\tilde{\boldsymbol{\theta}}')}{v_i(\theta_i') - v_i(\theta_i)} = g_j(\theta_j')A_{ji} + \sum_{k \neq i,j} g_k(\theta_k)\boldsymbol{A}_{ki}.$$

By subtracting the inequality when j has cost θ_j from the one when his cost is θ'_j we get

$$\frac{g_i(\theta_i)\left(x_i(\tilde{\boldsymbol{\theta}}) - x_i(\boldsymbol{\theta})\right) - g_i(\theta'_i)\left(x_i(\tilde{\boldsymbol{\theta}}') - x_i(\boldsymbol{\theta}')\right)}{v_i(\theta'_i) - v_i(\theta_i)} = \boldsymbol{A}_{ji}\left(g_j(\theta'_j) - g_j(\theta_j)\right).$$

However, note that $x_i(\tilde{\boldsymbol{\theta}}) - x_i(\boldsymbol{\theta}) = \boldsymbol{A}_{ij}\left(v_j(\theta'_j) - v_j(\theta_j)\right)$. Therefore,

$$\boldsymbol{A}_{ij}\frac{g_i(\theta_i) - g_i(\theta'_i)}{v_i(\theta'_i) - v_i(\theta_i)} = \boldsymbol{A}_{ji}\frac{g_j(\theta'_j) - g_j(\theta_j)}{v_j(\theta'_j) - v_j(\theta_j)}.$$

Recall that $\boldsymbol{A}$ is symmetric (Lemma B.4).Therefore, whenever $\boldsymbol{A}_{ij} \neq 0$ we must have:

$$\frac{g_i(\theta_i) - g_i(\theta'_i)}{v_i(\theta'_i) - v_i(\theta_i)} = \frac{g_j(\theta'_j) - g_j(\theta_j)}{v_j(\theta'_j) - v_j(\theta_j)}, \qquad \forall i \neq j,\ \forall \theta_i, \theta'_i \in \tilde{\Theta}_i,\ \forall \theta_j, \theta'_j \in \tilde{\Theta}_j. \tag{23}$$

Furthermore, the above equality should hold for every $i, j \in N$ as we can find a sequence of agents $\{l_0 = i, \ldots, l_K = j\}$ such that $\boldsymbol{A}_{l_k, l_{k+1}} \neq 0$ for all $0 \leq k < K$.[44]

To complete the proof that $g_i(\theta_i) = 0$ for all $\theta_i \in \tilde{\Theta}_i$, we are going to consider two options: either $g_i(\theta_i) - g_i(\theta'_i) = 0$ for at least one pair of $g_i(\theta_i), g_i(\theta'_i)$, or $g_i(\theta_i) - g_i(\theta'_i) \neq 0$ for all $i \in N$ and $\theta_i, \theta'_i \in \tilde{\Theta}_i$. We explore both options next:

Case $g_i(\theta_i) - g_i(\theta'_i) = 0$ for at least one pair of $g_i(\theta_i), g_i(\theta'_i)$. Suppose the numerator is zero for at least one pair of $g_i(\theta_i), g_i(\theta'_i)$. Then, $g_j(\theta_j) - g_j(\theta'_j)$ must be zero for every $j \in N$ and all pairs $\theta_j, \theta'_j \in \tilde{\Theta}_j$.

The next step is to show that $g_i(\theta_i) = g_j(\theta_j)$ must hold for every $\theta_i \in \tilde{\Theta}_i$ and $\theta_j \in \tilde{\Theta}_j$ and $i, j \in N$; we will use this fact as an intermediate step to show that indeed we must have $g_i(\theta_i) = 0$ for all $\theta_i \in \tilde{\Theta}_i$ and all $i \in N$.

Showing that $g_i(\theta_i) = g_j(\theta_j)$ is trivial if $i = j$, as $g_i(\theta_i) - g_i(\theta'_i)$ must be zero for every $i \in N$ and all pairs $\theta_i, \theta'_i \in \tilde{\Theta}_i$, by Eq. (23). For the other cases, note that when $g_i(\theta_i) = g_i(\theta'_i)$, we have $g_i(\theta_i)x_i(\boldsymbol{\theta}) - g_i(\theta'_i)x_i(\boldsymbol{\theta}') = g_i(\theta_i)\boldsymbol{A}_{ii}\left(v_i(\theta_i) - v_i(\theta'_i)\right)$. By Eq. (22) the above equality reduces to

$$\sum_{j \in N} g_j(\theta_j)\boldsymbol{A}_{ij} = 0, \tag{24}$$

and this must be true for any $i \in N$. Let $\boldsymbol{A}_R$ denote the submatrix of $\boldsymbol{A}$ consisting of $(n-1)$ linearly independent rows. By Lemma B.4, we know such matrix exists. Furthermore, we can

[44]Here we are implicitly assuming that matrix $\boldsymbol{A}$ has only one block. If $\boldsymbol{A}$ has more than one block, then we can use the same argument for each block.

assume that those are the $n-1$ demand equations that appear in the coefficient matrix $\boldsymbol{M}$. Let $\mathbf{g}=(g_1,\ldots,g_n)$ denote the vector of coefficients $g_i = g_i(\theta_i)$ for $\boldsymbol{\theta}\in\Theta$. By Eq. (24), the vector $\mathbf{g}$ must be in the nullspace of $\boldsymbol{A}_R$. However, as $\boldsymbol{A}_R\in\mathbb{R}^{(n-1)\times n}$ has dimension $(n-1)$ the dimension of its nullspace is at most 1. We will show that $\mathbf{1}$ is in Null$(\boldsymbol{A}_R)$, which implies that *all* g_i with $i\in N$ must be equal.

Consider $(\boldsymbol{A}_R)_{i,*}$, that is, row i of the coefficient matrix $\boldsymbol{A}_R$. We will show that $(\boldsymbol{A}_R)_{i,*}\cdot\mathbf{1}=0$. Note that

$$(\boldsymbol{A}_R)_{i,*}\cdot\mathbf{1}=\sum_j\left(-\boldsymbol{F}_{ij}+\frac{(\mathbf{1}'_{Q(\boldsymbol{\theta})}\cdot\boldsymbol{F}_{*,j})(\boldsymbol{F}_{i,*}\cdot\mathbf{1}_{Q(\boldsymbol{\theta})})}{\mathbf{1}'_{Q(\boldsymbol{\theta})}\boldsymbol{F}\mathbf{1}_{Q(\boldsymbol{\theta})}}\right)=-\boldsymbol{F}_{i,*}\cdot\mathbf{1}+\boldsymbol{F}_{i,*}\cdot\mathbf{1}=0,$$

as desired. Therefore, $\mathbf{1}$ is in Null$(\boldsymbol{A}_R)$ and thus $g_i(\theta_i)=g_j(\theta_j)$ for all $i,j\in N$, $\theta_i\in\tilde{\Theta}_i$, $\theta_j\in\tilde{\Theta}_j$.

Using that $g_i(\theta_i)=g_j(\theta_j)$ for all $\theta_i\in\tilde{\Theta}_i$ and $\theta_j\in\tilde{\Theta}_j$, we now show that $g_i(\theta_i)=0$ for all $i\in N$ and all $\theta_i\in\tilde{\Theta}_i$, which implies $b^i_{\theta_i}=0$ for all $\theta_i\in\tilde{\Theta}_i$. If $g_i(\theta_i)=0$, for some $i\in N$ and $\theta_i\in\Theta_i$, we are done. Otherwise, suppose that $g_i(\theta_i)=k\neq 0$ for all $i\in N$ and all $\theta_i\in\Theta_i$. By Lemma B.7 we have:

$$0\;=\;\sum_{j\in Q(\boldsymbol{\theta})}g_j(\theta_j)x_j(\boldsymbol{\theta})\;=\;k\left(\sum_{j\in Q(\boldsymbol{\theta})}x_j(\boldsymbol{\theta})\right)\;=k,$$

which is a contradiction ◁

Case $g_i(\theta_i)-g_i(\theta'_i)\neq 0$ for all $i\in N$ and $\theta_i,\theta'_i\in\tilde{\Theta}_i$. Let the pair $g_i(\theta_i),g_i(\theta'_i)$ be such that $\frac{g_i(\theta_i)-g_i(\theta'_i)}{v_i(\theta'_i)-v_i(\theta_i)}=k\neq 0$, and rewrite $g_i(\theta_i)=g_i(\theta'_i)+k[v_i(\theta'_i)-v_i(\theta_i)]$. Let $\theta_i,\theta'_i,\theta''_i\in\tilde{\Theta}_i$ and let $\theta_{-i}\in\tilde{\Theta}_{-i}$. Then, we must have using equation (22):

$$\begin{aligned}\big(v_i(\theta'_i)-v_i(\theta_i)\big)\sum_{j\neq i}\boldsymbol{A}_{ji}g_j(\theta_j)\;&=\;g_i(\theta_i)x_i(\boldsymbol{\theta})-g_i(\theta'_i)x_i(\boldsymbol{\theta}')\\&=\;\big(g_i(\theta'_i)+k[v_i(\theta'_i)-v_i(\theta_i)]\big)\,x_i(\boldsymbol{\theta})-g_i(\theta'_i)x_i(\boldsymbol{\theta}')\\&=\;g_i(\theta'_i)\big(x_i(\boldsymbol{\theta})-x_i(\boldsymbol{\theta}')\big)+k[v_i(\theta'_i)-v_i(\theta_i)]x_i(\boldsymbol{\theta})\\&=\;g_i(\theta'_i)\boldsymbol{A}_{ii}\big(v_i(\theta_i)-v_i(\theta'_i)\big)+k[v_i(\theta'_i)-v_i(\theta_i)]x_i(\boldsymbol{\theta})\end{aligned}$$

By dividing on both sides by $v_i(\theta'_i)-v_i(\theta)$ we obtain:

$$\sum_{j\neq i}\boldsymbol{A}_{ji}g_j(\theta_j)\;\;=\;\;-g_i(\theta'_i)\boldsymbol{A}_{ii}+kx_i(\boldsymbol{\theta})$$

In addition, since $\theta_i'' \in \tilde{\Theta}_i$, we have $\frac{g_i(\theta_i'')-g_i(\theta_i')}{v_i(\theta_i')-v_i(\theta_i'')} = k$ by Eq. (23), and thus:

$$\sum_{j \neq i} \boldsymbol{A}_{ji} g_j(\theta_j) \quad = \quad -g_i(\theta_i')\boldsymbol{A}_{ii} + kx_i(\boldsymbol{\theta}'')$$

which is a contradiction, because the virtual costs are strictly increasing and therefore $x_i(\boldsymbol{\theta}) \neq x_i(\boldsymbol{\theta}'')$ $\triangleleft$

Therefore, we have shown that in the first case we must have $g_i(\theta_i) = 0$ for all $\theta_i \in \tilde{\Theta}_i$ and all $i \in N$, and that the second case cannot arise as it will result in a contradiction. Note that this concludes the proof of Step 1.a —we have established that $g_i(\theta_i) = 0$ for all $\theta_i \in \tilde{\Theta}_i$ and all $i \in N$ $\triangleleft$

Step 1.b: $g_i(\theta_i) = 0$ **for all** $\theta_i \notin \tilde{\Theta}_i$ **and all** $i \in N$**.** Next, we show that $g_j(\theta_j) = 0$ whenever $\theta_j < \min \tilde{\Theta}_j$ or $\theta_j > \max \tilde{\Theta}_j$. (Recall that by Condition (2), $\tilde{\Theta}_j$ is an interval.) For $\theta_j < \min \tilde{\Theta}_j$ consider a profile $\boldsymbol{\theta} = (\theta_j, \boldsymbol{\theta}_{-j})$ such that $\theta_i \in \tilde{\Theta}_i$ for all $i \neq j$. By the definition of $\tilde{\Theta}_j$ and the monotonicity of demand (Lemma B.3), we must have $x_j(\boldsymbol{\theta}) > 0$. By Lemma B.7 and Step 1.a we have

$$0 \;=\; \sum_{i \in Q(\boldsymbol{\theta})} g_i(\theta_i) x_i(\boldsymbol{\theta}) \;=\; g_j(\theta_j) x_j(\boldsymbol{\theta}).$$

and therefore $g_j(\theta_j) = 0$ for all $\theta_j < \min \tilde{\Theta}_j$ and all $j \in N$.

Let $\theta_j > \max \tilde{\Theta}_j$ and $\theta_j \leq \theta_j^u$, as defined at the beginning of Section B.2. By Condition (5) of RIOM, there exists a profile $\boldsymbol{\theta} = (\theta_j, \theta_{-j})$ and a profile $\boldsymbol{\theta}' \in \tilde{\Theta}$ such that the profile $\boldsymbol{\theta}$ is reachable from $\boldsymbol{\theta}'$. Hence, there exists a sequence of profiles $\{\boldsymbol{\theta}_0 = \boldsymbol{\theta}', \ldots, \boldsymbol{\theta}_K = \boldsymbol{\theta}\}$ satisfying Condition (5). Given that $\boldsymbol{\theta}' \in \tilde{\Theta}$, we must have that $g_i(\theta_i') = 0$ for all $i \in N$. We will inductively show that $g_i((\boldsymbol{\theta}_k)_i) = 0$ for every $k = 1, \ldots, K$ and every $i \in Q(\boldsymbol{\theta}_k)$. As $j \in Q(\boldsymbol{\theta}_K)$, this will establish the result. Let k be the component in which $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ differ, and $k \in Q(\boldsymbol{\theta}_1)$. (If no such k exists, then all active agents share the same cost and thus the claim follows from the base case.) By Lemma B.7 we have $\sum_{i \in Q(\boldsymbol{\theta}_1)} g_i((\theta_1)_i) x_i(\boldsymbol{\theta}_1) = 0$. As $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ only differ in the k^{th} component, $\boldsymbol{\theta}_0 \in \tilde{\Theta}$, and $k \in Q(\boldsymbol{\theta}_1)$, we must have $g_k((\boldsymbol{\theta}_1)_k) = 0$. We can inductively repeat this argument to show that all the g's corresponding to a profile in the path between $\boldsymbol{\theta}'$ and $\boldsymbol{\theta}$ must be zero, which implies $g_j(\theta_j) = 0$ $\triangleleft$

Therefore, we have shown that both the statements described in Steps 1.a and 1.b hold. Therefore if $(\boldsymbol{a}, \boldsymbol{b})$ satisfies Eq. (12), then $g_i(\theta_j) = 0$ for all $i \in N$ and all $\theta_i \in \Theta_i$. This concludes the proof of Step 1. $\diamondsuit$

Step 2: $g_i(\theta_j) = 0$ for all $i \in N$ and all $\theta_i \in \Theta_i$ implies the system is consistent. So far we have shown that $g_i(\theta_j) = 0$ for all $i \in N$ and all $\theta_i \in \Theta_i$. By the definition of g_i, this implies $b^i_{\theta_i} = 0$ for all $i \in N$ and all $\theta_i \in \Theta_i$. To conclude the proof, we show that $b^i_{\theta_i} = 0$ for all $i \in N$ and all $\theta_i \in \Theta_i$ implies that the system is consistent. To that end, consider a vector $(\boldsymbol{a}, \boldsymbol{0})$ satisfying Eq. (12). For each $\boldsymbol{\theta} \in \tilde{\Theta}$, we have

$$\sum_{i=1}^{|Q(\boldsymbol{\theta})|-1} a^i_{\boldsymbol{\theta}} \left(\sum_{j \in Q(\boldsymbol{\theta})} \boldsymbol{A}_{ij}(\boldsymbol{\theta}) v_j(\theta_j) \right) = \sum_{j \in Q(\boldsymbol{\theta})} v_j(\theta_j) \left(\sum_{i=1}^{|Q(\boldsymbol{\theta})|-1} a^i_{\boldsymbol{\theta}} \boldsymbol{A}_{ij}(\boldsymbol{\theta}) \right) = 0,$$

as $(\boldsymbol{a}, \boldsymbol{0})$ satisfying Eq. (12) implies $\sum_{i=1}^{|Q(\boldsymbol{\theta})|-1} a^i_{\boldsymbol{\theta}} \boldsymbol{A}_{ij}(\boldsymbol{\theta}) = 0$. Hence, we have shown that $(\boldsymbol{a}, \boldsymbol{0})$ also satisfies Eq. (13). Therefore, the system is consistent and $OPT(P_1) = OPT(P_0)$ as desired. $\square$