# Competition in Service Operations and Supply Chains: Equilibrium Analysis and Structural Estimation

**Lijian Lu**

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

**COLUMBIA UNIVERSITY**

2016

# ABSTRACT

# Competition in Service Operations and Supply Chains: Equilibrium Analysis and Structural Estimation


## Lijian Lu



The service industry has become increasingly competitive. This dissertation addresses a number of outstanding and fundamental questions of competitions in service operations and supply chains. The challenges are characterization of the equilibrium behaviors, estimating the impact of firms' interactions, and designing of efficient market mechanisms.

The first chapter of this dissertation considers price competition models for oligopolistic markets, in which the consumer reacts to *relative* rather than *absolute* prices, where the relative price is defined as the difference between the absolute price and a given reference value. Such settings arise, for example, when the full retail price earned by the "retailer" is reduced by virtue of a third party offering a subsidy or a rebate or in prospect theoretical models in which customers establish a reference price and base their choices on the differentials with respect to the reference price. When choosing among the various competing options, the consumer trades off the net price paid with various other product or service attributes, as in standard price competition models. The reference price may be *exogenously* specified and pre-announced to the competing firms. Alternatively, it may be *endogenously* determined, as a function of the set of absolute prices selected by the competing firms, for example the lowest or the second lowest price. We characterize the equilibrium behavior under a general reference value scheme of the above type; this in a base model, where we assume that the consumer choice model is of the general MultiNomialLogit (MNL) type. We also derive comparison results for the price equilibria that arise under alternative subsidy schemes. These comparisons have important implications for the design of subsidy schemes.

The second chapter applies the results of the first chapter to the Medicare insurance market, both in terms of its existing structure, as well as in terms of various proposals to redesign the program. Based on an oligopoly price competition model tailored towards this market, and actual

county-by-county data for the year 2010, we estimate the impact such reforms would have on the plans' market shares, equilibrium premia, the government's cost, and the out-of-pocket expenses of the beneficiaries. We employ two different methodologies to derive the parameters in the county-by-county competition models: (i) a calibration model, and (ii) parameter distributions obtained from models estimated in Curto *et al.* [2015]. The predicted impacts on the above performance measures are remarkably consistent across the two methodologies and reveal, for example, that the government cost would decrease by 8% if the traditional fee-for-service(FFS) plans are kept out of competitive bidding process and by 16.5%-21% if they are part of the process.

The third chapter studies a class of buy procurement mechanisms, framework agreements (FAs), that are commonly used by buying agencies around the world to satisfy demand that arises over a certain time horizon. We are one of the first in the literature that provides a formal understanding of FAs, with a particular focus on the cost uncertainty faced by bidders over the FA time horizon. We introduce a model that generalizes standard auction models to include this salient feature of FAs; we analyze this model theoretically and numerically. First, we show that FAs are subject to a sort of winner's curse that in equilibrium induces higher expected buying prices relative to running first-price auctions as needs arise. Then, our results provide concrete design recommendations that alleviate this issue and decrease buying prices in FAs, highlighting the importance of (i) monitoring the price charged at the open market by the FA winner and using it to bound the buying price; (ii) investing in implementing price indexes for the random part of suppliers' costs; and (iii) allowing suppliers the flexibility to reduce their prices to compete with the open market throughout the selling period. These prescriptions are already being used by the Chilean government procurement agency that buys US$2 billion worth of contracts every year using FAs.

The fourth chapter considers the preference of contractual forms in supply chains. The supply chain contracting literature has focused on incentive contracts designed to align supply chain members' individual interests. A key finding of this literature is that members' preferences for contractual forms are often at odds: the upstream supplier prefers more complex contracts that can coordinate the supply chain; however, the downstream retailer prefers the wholesale price–only contract because it leaves more surplus (than a coordinating contract) that the retailer can get. This chapter addresses the following question: under what circumstances do suppliers and retailers prefer the same contractual form? We study supply chain members' preference for contractual forms in three different competitive settings in which multiple supply chains compete to sell substitutable products to the same market. Our analysis suggests that both upstream and downstream sides of the supply chains may prefer the same "quantity discount" contract, thereby eliminating the conflicts of interest that otherwise typify contracting situations. More interesting still is that both sides may also prefer the wholesale price–only contract, which offers a theoretical explanation to why the simple inefficient contract is widely adopted in supply chain transactions.

# Contents

# List of Figures

# List of Tables

# Acknowledgments

My sincerest and deepest gratitude to my advisors Professor Awi Federgruen and Professor Gabriel Weintraub. It is my great honor to work with them. I would not have accomplished this without their continuous guidance and encouragement throughout the years.

I own my heartfelt appreciation to Professor Omar Besbes, Fanyin Zheng, and Jay Sethuraman for serving on my defense committee. I am also grateful to Professor Fangruo Chen for his advices and supports throughout my doctoral life. I would also like to thank Professor Carri Chan, Linda Green, Jeannette Song, Yaozhong Wu, and Hanqin Zhang for sharing thought and experience about research and academic life in general.

Last but not the least, I would like to give my special thank to my wife Ji Wen, and my parents Shixing Lu and Yuying Rao. Their love and support throughout these years have been a source of joy and a pillar of strength for me to overcome hard times.

To Ji, Shixing, and Yuying

# Chapter 1

# Price Competition Based on Relative Prices

Awi Federgruen and Lijian Lu

## 1.1 Introduction and Summary

Ever since the seminal paper by Bertrand [1883], a vast literature has developed on price competition models in oligopolistic markets, see e.g. Tirole [1988], Topkis [1998] and Vives [2001]. Almost invariably, these models assume that customers choosing among various products, react to the absolute (unit) prices charged by the competing firms, hereafter referred to as the "retailers". There are, however, many contexts in which the consumer reacts to *relative* rather than *absolute* prices, where the relative price is defined as the difference between the absolute price and a given reference value.

Such settings arise, for example, when the full retail price earned by the "retailer" is reduced by virtue of a third party offering a subsidy or a rebate, e.g., via a coupon. The subsidy or rebate acts as the reference value and the consumer only relates to the relative prices of the competing products. The third party may be a government agency or international organization interested in promoting the adoption of the product or technology, e.g., solar panels, hybrid cars, electric cars, vaccines, or because they consider the full cost of an essential product or service prohibitive, e.g., health insurance, as in the US Medicare program. Alternatively, the third party may consist of manufacturers offering a subsidy or coupon program for their product(s) to stimulate retail sales, or insurance companies covering a large part of the product or service cost.( Pharmaceutical products are an example; here, the insured consumer only pays the so-called co-payment.)

At the same time, empirical studies and laboratory experiments have, consistently, demonstrated that consumer choices often defy the predictions of standard rational choice models. The most important breakthrough in the field of behavioral economics is "Prospect Theory", developed by Kahneman and Tversky [1979] and Tversky and Kahneman [1991]. The principal idea of prospect theory is that individuals evaluate different alternatives or outcomes, based on the differentials with respect to some objective or subjective reference point, rather than the absolute values of the relevant attributes. In terms of price comparisons, this implies that consumers establish a "reference price" and that their utility measures are explained by the relative price differentials of the various products with respect to the reference price.

Within the Marketing literature, this paradigm was first addressed by Winer [1986]. He proposed a MultiNomial Logit (MNL) model, in which the relative price appears as a (linear) term in the specification of each brand's utility measure. The author proceeded to fit this model to data for a retail coffee market with three competing brands, estimating that the coefficient of the relative price is significant in at least two of the three brands' utility equations. Winer [1986] entertained multiple structures for the determination of the relative price.

Subsequent to Winer's seminal paper, at least 16 similar models were proposed in the premier Marketing journals, in the subsequent 20 years, as reviewed by Mazumdar *et al.* [2005], building on an earlier survey article by Briesch *et al.* [1997]. Some follow Winer's specification, in that the relative price appears as a linear term in the products' utility measures.(Mazumdar *et al.* [2005] refer to this structure as the "*symmetric sticker shock model*".) Others recognize, that a positive relative price causes a larger reduction in the product's utility, as compared to the increase in this measure due to an equally large negative relative price, introducing a piecewise linear dependence of the utility measure on the relative price.(Mazumdar *et al.* [2005] refer to this generalization as the "*asymmetric reference price model*".) Empirical evidence consistently suggests that "losses" are weighted more heaviliy than equally sized "gains", a paradigm called "loss aversion", see Tversky and Kahneman [1991].

The main features of the above 16 MNL models are summarized in Table 2 of Mazumdar *et al.* [2005]. The models were used to explain brand choices in various food industries, as well as the liquid detergent industry, see Mazumdar and Papatla [2000] and the detergent, paper towel and tissue industries in Bell and Lattin [2000]. The impact of the reference value, and hence the relative, as opposed to the absolute price was confirmed by the estimation results, in virtually all of the 16 studies.

Similarly, in Section 1.4, we review a variety of industries in which the retail price is subsidized by a third party, whether a government agency or a private company, so that the consumer's actual expense is given by a relative price. These include the Medicare insurance and various green technology industries with very sizable government subsidies and industries in which manufacturer coupons are offered to the consumer. As with the above behavioral marketing models, see equation (5)-(7) in Mazumdar *et al.* [2005], almost all of the proposed models are of the MNL type.

Thus, the MNL model with utility measures that depend on a relative price, has become the workhorse model in many settings, and has the potential to be used as such, in various other applications. However, in spite of its 30 year history, little is known about its competitive dynamics, in particular its equilibrium behavior, specifically the following questions:

(I) Is a pure Nash equilibrium guaranteed to exist?

(II) Assuming it exists, is (are) the equilibrium (equilibria) globally stable, in the sense that, if the industry starts at an arbitrary price vector and firms iteratively adjust their prices in response to their competitors' price choices, this process will converge to an equilibrium?

(III) Assuming a pure Nash equilibrium exists, is such an equilibrium unique, and if not ,what can be said about the structure of the set of equilibria?

(IV) Assuming an equilibrium is guaranteed to exist, how does the equilibrium or the set of equilibria, vary as the reference value is altered?

As explained in more detail, below, a change in the reference value may involve a simple shift in its level, as when an exogenously specified subsidy or coupon amount is increased or decreased. Alternatively, the change may be structural, when the very structure of the reference value's dependence on the competitors' prices is modified.

The objective of this paper is to resolve all of the above questions (I)-(IV) for a general MNL model in which the products' utility measure is a monotone, generally non-linear, response function of its relative price, and this under various prevalent structures for the reference value.

The above questions are not just of theoretical importance. They are, often, at the heart of private or public policy debates. Consider, for example, the above Medicare industry, discussed in more detail in Section 3.1: private insurance companies offer plans to eligible consumers, as an alternative to the traditional (fee-for-service) Medicare plan offered by the federal government. In the current system, the government announces, every year, a (county specific) capitation rate or subsidy, before the private insurance firms submit their plans and associated prices (premia). There are continuous proposals to alter the capitation levels, some of which have been legislated, for example in the 2000 Benefits Improvement and Protection Act (BIPA) and the 2011 Affordable Care Act (ACA). Policy makers struggle to predict what impact these level changes would have on the new equilibrium premia, the firms' market shares, the consumers' out of pocket costs, government expenditures etc. Several prominent recent studies, e.g., Song *et al.* [2012b], Song *et al.* [2013], Cabral *et al.* [2014] and Duggan *et al.* [2014] have applied regression models to estimate the pass-through rates, i.e., the absolute change in the premia due to changes in the capitation rates. These studies reach rather different conclusions, with estimated pass through rates varying between

37% and 100% across the different studies, leaving policy makers with a great deal of ambiguity. Moreover, these reduced form approaches assume a specific structural form, for example linear or log-linear, for the dependency of the firms' premia on the subsidy (capitation) levels, which may not be consistent with any plausible underlying price competition models.

For example, based on the results of this paper, Federgruen and Lu [2016b] *calculate* pass through rates in an MNL-model calibrated to the actual county-by-county data in the 2010 Medicare market. Their results show: (1) the pass through rate fails to be constant, neither in absolute nor in relative terms, as implied, respectively, by a linear or log-linear regression model. Instead, as the capitation rate is reduced by up to $30 from their prevailing average ($802), the pass through rate varies in a close to a 3:1 ratio; (2) different competing plans adopt rather disparate pass through rates, and (3) the estimated (average) pass through rates are considerably *lower* than those estimated from the above regression equations, see Figure 1.1 in Section 1.4.

Finally, even if the pass through rates are estimated correctly, to assess the various above performance measures, there is a need to predict how the changes would affect the firms' market shares. To this end, Curto *et al.* [2015] develop and estimate an MNL-like consumer choice model[1] for the Medicare industry, along with a regression model to predict pass through rates. These models are then applied to conduct several counterfactual studies. Among them, is an assessment of an across the board $50 reduction in the monthly capitation rate. The regression equations are used to estimate the new premia bids; these are then entered into their consumer choice model to evaluate the firms' new market shares. It seems preferable to conduct the entire counterfactual study within a single consistent estimated consumer choice model, but this requires a resolution of all of the equilibrium behavior questions (I)-(IV), raised above, including the ability to compute the new price equilibrium.

Moreover, reform proposals for the Medicare program have by no means been restricted to level changes in the benchmark or subsidy level. Instead, there have been proposals to determine each county's subsidy to the private plans, endogenously, as either the second lowest or a weighted average of the bids. These ideas were the core of proposals by The Bipartisan Policy Center's Debt Reduction Task Force, various House passed Budget resolutions, the 2012 bipartisan reform plan

[1]More precisely, the model is an application of the so-called nested MNL model with all private plans in a single group and the traditional government plan , by itself, in a second group.

initiated by Senator Wyden and Congressman Ryan, and, most recently, the proposals in Rivlin and Daniel [2015], as part of an NBER conference, entitled “Strengthening Medicare for 2030”. A variety of studies have tried to examine the implications of these endogenously determined capitation values. For example, Song *et al.* [2012a] and more recent estimates by the Congressional Budget Office (2013) concluded that the average capitation rate, and hence the government’s overall expenditure, would be reduced by 9% and 11% respectively, were it to be determined by the second lowest bid among the private plans.

However, these estimates are based on the assumption that all plans would continue to be offered at the same premium and attract the same market share. In reality, firms, of course, adjust their prices to the new rules, also resulting in a new set of market shares. But to assess these price and market share changes requires an (MNL-like) consumer choice model, as in Federgruen and Lu [2016b], Einav and Levin [2015], and Curto *et al.* [2015] and a full resolution of the above questions regarding the model’s equilibrium behavior under various mechanisms to determine the reference value or capitation rate. Indeed, based on the results in this paper, Federgruen and Lu [2016b] are able to show that firms would reduce their premium bids, significantly, resulting in an additional decrease of the capitation rates and an effective decrease of government expenditures by 16% (, as opposed to 11%)!

### Our main results

As mentioned, we resolve the above questions (I)-(IV) for general MNL models with the products’ utility measures specified by a general monotone response function of the relative price, und this under the following 5 reference value structures:

(R1) an exogenously given constant,

(R2) an exogenously given fraction of the absolute price,

(R3) the lowest price in the market,

(R4) the n-th lowest price,

(R5) a (weighted) average of the selected market prices.

(This spectrum covers most applications we are aware of, including the above mentioned Marketing papers with reference pricing, see Mazumdar *et al.* [2005].) See Cadotte *et al.* [1987] and Köszegi and Rabin [2006] for an in-depth discussion and derivation of reference value structures.

The importance of considering a general monotone response function is driven by behavioral models, as in the above asymmetric reference price models, or even more general response functions suggested by prospect theory. They are also motivated by concrete subsidy structures as in the Medicare market and ACA exchanges, discussed in section 1.4.1.

As to the first question (I), we show that a pure Nash equilibrium exists under reference value structures (R1)-(R4), and very general conditions for the response function. Under (R5), i.e., when the reference value is specified as the average price (or, more generally, a weighted average of the selected prices), we show that the products fail to be substitutes, i.e., a price increase of one product may result in a *decrease* of the demand volume of one of its competitors. Since the products are no longer substitutes, the existence of a Nash equilibrium can not be guaranteed, a well known phenomenon, see e.g. Vives [2001] Section 2.3.2.

Global stability, i.e., an affirmative answer to question (II) is proven, under mild conditions, for reference value structures (R1)-(R3), arguably the most important cases from a practical perspective. (Given the fact that a pure Nash equilibrium cannot be guaranteed under (R5), question (II) is moot under this structure.) While a strong form of equilibrium, global stability, ensures that the set of equilibria has a componentwise smallest and a componentwise largest element, and a lattice structure. It also provides a simple algorithm to compute an equilibrium using a tâtonnement scheme. Moreover, the scheme can be used to verify whether the equilibrium is unique, thus answering question (III).

To address question (IV), it is useful to differentiate between the following two cases, both of which have ample applications, see Section 1.4.

(I) (PIOG) *Price Invariant Outside Good*: here no price choice is made for the outside good and its utility measure is independent of the reference value.

(II) (PSOG) *Price Sensitive Outside Good*: here the utility measure of the outside good depends on the price and the reference value, via the same response function as the other goods.

(I) (*PIOG*): Under (R1), we prove that the equilibrium prices are monotonically increasing in the exogenous, constant reference value, under a mild condition for the outside good's market share. Similarly, we show, under (R2), that the equilibrium prices are monotonically increasing in the discount percentages specified by the reference value, provided the latter is sufficiently large. When comparing the equilibria under (R1) with those under (R2), we establish a threshold result for a widely applicable class of response functions: that includes the above "symmetric sticker shock" and the "asymmetric reference price" models: fix the constant exogenous reference value under (R1); for any given product, the equilibrium price under (R1) is exceeded by that under (R2), if and only if, the discount percentage is larger than the threshold. A similar threshold result applies, when the equilibrium under (R1) is compared to that under (R3), where the reference value is determined as the *lowest* price. All of the above results apply both to the component-wise smallest and the component-wise largest equilibrium.

(II) (*PSOG*): Our numerical experiments show that, *in general*, monotonicity of the price equilibrium vector in an exogenous constant reference value continues to hold. However, monotonicity is sometimes violated, as a counterexample demonstrates. In contrast, under (R2), we are able to prove that monotonicity of the equilibrium price with respect to the discount percentages continues to hold, similar to this monotonicity result under (PIOG). When comparing (R1) and (R2), a similar threshold result as under (PIOG) continues to prevail. When comparing (R1) and (R3), i.e., an exogenous constant reference value vis-a-vis one determined by the lowest price, we prove, under a mild condition, that the equilibrium under (R3) is *always* exceeded by that under (R1).

The remainder of this paper is organized as follows: Section 1.2 provides a literature review. In Section 2.4, we present the general model. Section 1.4 contains a brief discussion of several applications. Section 1.5 characterizes the equilibrium behavior when the reference value is an exogenous constant or a given fraction of the absolute prices. Section 1.6 covers the case where the reference value is endogenously determined as the lowest price, or more generally the n-th lowest price. The above discussed comparison results of the price equilibria under various reference value schemes are derived in Section 1.7. Section 2.7 concludes our paper and discusses several extensions. Most of the proofs are deferred to the Appendix.

## 1.2 Literature Review

Several papers have addressed the equilibrium behavior in MNL-type models, but all under the assumption that the utility measures' price dependence is confined to a dependence on the the product's own and absolute price. It is well known that an equilibrium exists, in the standard MNL models, see Anderson *et al.* [2001], Bernstein and Federgruen [2004] and Gallego *et al.* [2006]. However, an equilibrium may fail to exist in various generalizations of the basic MNL model, for example Mixed MNL models (MMNL) where the market is segmented and the structure of the utility functions varies by segment.(In the latter case, Allon *et al.* [2013] have shown that an equilibrium may fail to exist while providing specific market share conditions under which the existence question can be answered in the affirmative.) Liu [2006], Li and Huh [2011], and Gallego and Wang [2014] study various conditions under which an equilibrium is guaranteed to exist in a nested logit model.

To our knowledge, our paper provides the first characterizations of the equilibrium behavior of a price competition model based on a MNL consumer choice model, in which utilities depend on relative prices. Heidhues and Kőszegi [2008] recently analyzed a Hotelling type price competition model in which the utility measure of each product depends on its price via a piecewise linear function of the *difference* between the price and an exogenous reference value, i.e., the (R1) structure. In this model, customers are uniformly dispersed on a unit circle and $N$ firms are located on this circle as well. All heterogeneity among customer preferences is represented by a "distance cost" term in their utility measures, assumed proportional to the distance (along the circle) between the customer's location and that of the selected firm. The firms face a firm-specific stochastic cost rate. The authors provide conditions under which a symmetric price equilibrium exists and other conditions under which every equilibrium is symmetric. Several authors, in particular, Zhou [2011] and Karle and Peitz [2014] have characterized the equilibrium behavior in variants of the Heidhues and Kőszegi [2008] model, but all for a special case of a *duopoly.* See Ellison [2006] and Spiegler [2011] for discussions of the importance of price competition models in which consumers are assumed to be loss averse relative to a reference value.

A similar prospect theoretical competition model has recently been proposed by Yang *et al.* [2014], to model competition in service industries. Here, firms select and announce waiting time standards, for example the steady-state average waiting time. However, individual customers' utility depends on the *relative* waiting time defined as the *difference* between the actual waiting time and a reference value. The authors address the case where the reference value is given by the *lowest* waiting time standard, as in (R3), as well as where it is specified as a weighted average thereof, as in (R5). Among others, the following are four important differences between the Yang *et al.* [2014] model and ours. First, in an MNL model, like ours, heterogeneity among the customers is expressed via the unobserved noise terms in the utility measures. In Yang *et al.* [2014], similar to Heidhues and Kőszegi [2008], all heterogeneity among customer preferences is confined to a distance cost term in the utility measure, assumed to be proportional with the distance between the customer's initial position and the location of the chosen firm, where, in this model, the former is uniformly distributed on a symmetric hub-and-spoke network. Second, Yang *et al.* [2014] is confined to a duopoly, i.e., $N = 2$. Third, their model assumes that the two firms have identical characteristics, with the exception of their waiting time standards, and fourth, it assumes that each customer has to select one of the two providers, without a no-purchase or outside option.

Recently, several areas in operations and revenue management have come to realize that demand volumes should be represented as functions of *relative* prices, i.e., absolute or relative differentials of the nominal price vis-a-vis a reference value. For example, Chen *et al.* [2016]'s dynamic pricing model, represents the demand volume as an *asymmetric reference price model*, similar to Marketing papers surveyed in Mazumdar *et al.* [2005]. Similarly, Johnson-Ferreira *et al.* [2016] present a price optimization model developed for an online retailer, Rue La La, in which the demand for an item is represented as a function of the ratio of item's price and a reference value, defined as the *average* price across all competing items in the product category, see (R5).

## 1.3 Model

Consider an oligopolistic market with $N$ competing single-product firms each selling a product or service. The firms differentiate themselves via an arbitrary collection of observable product characteristics, as well as their nominal or *absolute* price. Potential customers react to *relative* prices when comparing the various product alternatives, either because of third party subsidies or because of prospect theoretical considerations, see the Introduction. A product's relative price is defined as the difference between the absolute price and a reference value. The reference value (i.e., subsidy or prospect theoretical benchmark value) may be an *exogenously* specified constant, which is pre-announced to the competing firms. Alternatively, it may be *endogenously* determined, as a function of the set of nominal prices selected by the competitors, for example the lowest, the second lowest, or an average price, see the reference values (R1)-(R5). Each firm's cost structure is assumed to be affine. For each firm $i = 1, 2, \dots, N$, let

$$\begin{aligned}
c_i &= \text{the marginal cost rate of providing product } i \\
p_i &= \text{nominal price of the product or service provided by firm } i\text{, to be selected from an} \\
&\quad\ \text{interval } [p_i^{min}, p_i^{max}] \\
p_{-i} &= \text{price vector for all firms other than } i\text{, i.e., } p_{-i} = (p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_N) \\
p_{(n)}^{-i} &= \text{the } n\text{th smallest price excluding } p_i \text{ if } n \geq 1\text{, and } c_i \text{ if } n = 0 \\
p_{(n)} &= \text{the } n\text{th smallest price, } n \geq 1 \\
g_i(\mathbf{p}) &= \text{the reference value for product } i \\
\Delta p_i &= \textit{net} \text{ price of product } i = p_i - g_i(p) \\
d_i &= \text{market share of firm } i
\end{aligned}$$

We choose $p_i^{min} = c_i, i = 1, 2, \dots, N$. Based on the structures encountered in various applications, we consider the following specifications of the reference values, corresponding with structures (R1)–(R5) in the Introduction:

$$g_i(\mathbf{p}) = \begin{cases} C, & \text{exogenous and constant (R1)} \quad (1.1a) \\ \delta_i p_i, & \text{a percentage discount of the nominal price (R2)} \quad (1.1b) \\ p_{(1)}, & \text{specified by the lowest price (R3)} \quad (1.1c) \\ p_{(n)}, & \text{specified by the } n\text{-th lowest price (R4)} \quad (1.1d) \\ \sum_j w_j p_j, & \text{a weighted average of selected prices (R5)} \quad (1.1e) \end{cases}$$

(In most of the above specifications, all products share an identical reference value, i.e., $g_1(\mathbf{p}) = g_2(\mathbf{p}) = \cdots = g_N(\mathbf{p})$. However, under (R2), for example, when a third party (e.g., a manufacturer) offers a given percentage discount as a subsidy, the discount percentage $\delta_i$ may be product specific.) The reference structures (R1)-(R2) are studied in Section 1.5 and (R3)-(R4) in Section 1.6. Structure (R5) is discussed as part of Section 2.7.

Each customer $j$ assigns a utility measure to each of the $N$ available products, as follows

$$u_{ij} = a_i - b_i \cdot f\left(p_i - g_i(\mathbf{p})\right) + \epsilon_{ij}, \quad i = 1, 2, \ldots, N, \ j = 1, 2, \ldots. \tag{1.2}$$

Here, the intercept $a_i$ denotes the aggregate impact of all of the product's observable attributes, with the exception of the price. The, generally non-linear, function $f(\cdot)$ characterizes how the relative price impacts on the utility measure. A non-linear and not necessarily differentiable choice for this function may often be necessary; for example, in the Medicare industry, if the net price is negative, the customer receives less than (the absolute value of) this net price as a rebate, in accordance with a specific rebate percentage. In addition, the marginal disutility due to an extra \$10 of out-of-pocket expenses, may not be constant. Likewise, in prospect-theoretical models, it is well known that changes from the reference value are weighted differently, depending upon whether they are gains or losses, again requiring a non-linear response function $f(\cdot)$. We allow for the price sensitivity coefficient $b_i$ to be product specific. Finally, the last term $\epsilon_{ij}$ in (2.2) represents a random unobserved component of the customer $j$'s utility for product $i$, which varies by customer.

We represent the outside good as a product 0, with a utility measure

$$u_{0j} = a_0 - b_0 f\left(p_0 - g_0(\mathbf{p})\right) + \epsilon_{0j}, \quad j = 1, 2, \ldots, \tag{1.3}$$

where $a_0$ is a constant and $\epsilon_{0j}$ is a random unobserved component.

*Price Invariant Outside Good (PIOG)*: In some applications, consumers may choose not to purchase any variant of the product, in which case the utility measure is best described without the second term in (2.4), i.e., selecting $b_0 = 0$. In other applications, e.g., the Medicare market, discussed in the Introduction and Section 1.4.1, all Medicare beneficiaries enroll in one of the plans, but we use product 0 to refer to the traditional Medicare plan, whose enrollees currently pay nothing beyond the basic premium charged to all beneficiaries. Here, too, a specification with $b_0 = 0$ is called for.

*Price Sensitive Outside Good (PSOG)*: However, in some of the the Medicare reform proposals, those enrolled in the traditional Medicare program would be charged on the basis of the net price, relative to the same capitation rate that applies to the MA plans. To model these settings, a specification with $b_0 = b > 0$ is required.

To complete the specification of the utility functions (2.2), the random variables $\{\epsilon_{ij}\}$ for the unobserved utility components are assumed to be i.i.d across firms and customers, following a standard type 1–extreme value or Gumbel distribution, i.e., $\mathbb{P}(\epsilon_{ij} \leq x) = \exp(-\exp(-x+\gamma))$ where $\gamma$ is Euler's constant (0.5772).(The mean and variance of $\epsilon_{ij}$ are $E[\epsilon_{ij}] = 0$ and $var[\epsilon_{ij}] = \pi^2/6$.) This gives rise to a variant of the famous MNL model, with the following expected demand functions:

$$d_i(\mathbf{p}) = \frac{\exp(a_i - b_i \cdot f(p_i - g_i(\mathbf{p})))}{\exp(a_0 - b_0 f(p_0 - g_0(\mathbf{p}))) + \sum_{j=1}^{N} \exp(a_j - b_j \cdot f(p_j - g_j(\mathbf{p})))}, i = 0, 1, \ldots, N, \quad (1.4)$$

see for e.g., Anderson *et al.* [2001]. The above specification treats the utilities of all customers as identically distributed. Often, the market needs to be segmented into several customer classes, each with its own specification of the utility measures. This generalization is referred to as a Mixed-MultiNomialLogit (MMNL) model, briefly described in Section 2.7.

The response function is taken as a general increasing and usually non-linear function, subject to a minor regularity condition:

**Assumption 1** $f(\cdot)$ *is strictly increasing and continuous everywhere; it is continuously differentiable everywhere, with the possible exception of a finite set* $\mathbb{P}$*, where the function has a right- and left-hand derivative.*

Let

$$\alpha = \max_{i=0,1,\ldots,N} \sup\{f'(p_i - g_i(\mathbf{p})) : \ p_j \in [p_j^{min}, p_j^{max}] \ \forall j \text{ and } p_i - g_i(\mathbf{p})) \notin \mathbb{P}\},$$
$$\beta = \min_{i=0,1,\ldots,N} \inf\{f'(p_i - g_i(\mathbf{p})) : \ p_j \in [p_j^{min}, p_j^{max}] \ \forall j \text{ and } p_i - g_i(\mathbf{p})) \notin \mathbb{P}\}.$$

**Lemma 1** $0 < \beta \leq \alpha < \infty$.

We refer to the ratio $\beta/\alpha \in [0, 1]$ as the *degree of non-linearity* of the response function. This index plays a fundamental rule in many of our results. For some of our results, stronger conditions are required for the response function, in particular convexity. Convexity holds in many of the applications, in particular the above mentioned asymmetric reference price model adopted in the Marketing literature, see Mazumdar *et al.* [2005]. It also applies in the Medicare market described in Section 1.4.1, as well as the base model in prospect theory. However, some prospect theoretical models, employ an S-shaped response function which fails to be convex on its full domain.

Each firm earns its selected nominal price, under all circumstances, for each of its customers, and its profit function is given by

$$\pi_i(p_i, p_{-i}) = (p_i - c_i)d_i(\mathbf{p}). \tag{1.5}$$

Similarly, the aggregate consumer surplus is given by

$$CS(\mathbf{p}) = -\sum_i d_i(\mathbf{p})f(p_i - g_i(\mathbf{p})). \tag{1.6}$$

For any function $H(x)$, we denote the left-limit and right-limit at a particular point $x_0$ by $H_-(x_0) \equiv \lim_{x \nearrow x_0} H(x)$ and $H_+(x_0) \equiv \lim_{x \searrow x_0} H(x)$, respectively. Similarly, we write $\frac{\partial_- H}{\partial x}(x_0) = \lim_{u \nearrow x_0} \frac{\partial H}{\partial x}(u)$ and $\frac{\partial_+ H}{\partial x}(x_0) = \lim_{u \searrow x_0} \frac{\partial H}{\partial x}(u)$, respectively, whenever these limits exist. For any real number $x$, $x^+$ and $x^-$ denote the positive and negative part of $x$.

## 1.4 Applications

In this section, we provide a brief description of several industries to which the competition model of Section 2.4 is applicable.

### 1.4.1 The Medicare Advantage Market

In the US, all citizens and permanent residents, 65 years or older, are eligible to enroll in the Medicare program. In 2014, the Medicare program, covered approximately 54 million individuals, at an annual cost close to half a trillion dollars, or 14% of total federal spending. Moreover, without any restructuring, Medicare costs are estimated to grow at twice the rate of the GDP, the result of the upcoming retirement of many baby boomers, increased longevity, as well as the escalating costs of healthcare. The Congressional Budget Office (CBO) has estimated that the government's healthcare liabilities, as a percentage of the GDP, would grow from 5% to 12% in the next 40 years, a situation most politicians and economists consider untenable, see e.g., Rivlin and Daniel [2015].

An individual who is eligible for Medicare coverage, has the choice of enrolling in the traditional government plan or one of the approved private plans, since 2003, referred to as the Medicare Advantage (MA) plans. In 2003, private MA plans captured only 13% of the potential market; however, their share has steadily grown to 30% in 2014.

As mentioned, in the MA program, the government pays most of the insurance premium of the different plans that are offered to the beneficiaries. Insurance companies submit, each year, by a given deadline, one or several plans, each covering one or a collection of counties. In the current structure, the government announces a county-specific capitation rate or premium. All individuals covered by Medicare pay a monthly base premium, irrespective of their plan choice, When choosing a private plan, a beneficiary pays an additional premium given by the relative premium (= nominal premium - capitation rate), when positive, or receives a 75% rebate when the relative premium is negative.

Among other such proposals, the 2011 Wyden-Ryan (W-R) plan, prominently debated in the 2012 presidential campaign, advocates replacing the exogenous capitation rate by the second-lowest bid. More recently, the Congressional Budget Office [2013] has considered, as an alternative to the second lowest bid scheme, one which specifies the subsidy as a weighted average of the premium bids (with the past year's market shares as the weights). The report estimates that the second lowest bid option [average bid option] results in an 11% [4%] reduction of federal spending, while increasing payments by affected beneficiaries by 11% [-6%]. As with prior such estimates, e.g., Song *et al.* [2012a], they are based on the simplifying assumption that the premium bids and market shares would not be affected by the change in the subsidy scheme, when, in reality, of course, they would. To assess the changes in the bid prices and market shares, one needs to model the competition among the insurance companies as a price competition model with consumer utility measures dependent on relative prices, for example, the model in Section 2.4.

To fit this model to the MA market in a given county, one needs to specify $a_i$ as a function of the plans' non-premium attributes. For example, Curto *et al.* [2015] consider indicator variables describing whether a plan has a given quality standard, and whether it is associated with several supplemental benefits, e.g., vision and dental coverage. To capture the above asymmetric consequences of positive versus negative relative prices, the response function should be specified as

$$f(x) = x^+ - 0.75x^- \tag{1.7}$$

a convex piece-wise linear function, of type (1.9). Depending upon whether the capitation rate is exogenously announced, or endogenously determined as the second lowest bid or a weighted average of the bids, we specify:

$$g(\mathbf{p}) = \begin{cases} C \\ p_{(2)} \\ \sum_i w_i p_i \end{cases} \tag{1.8}$$

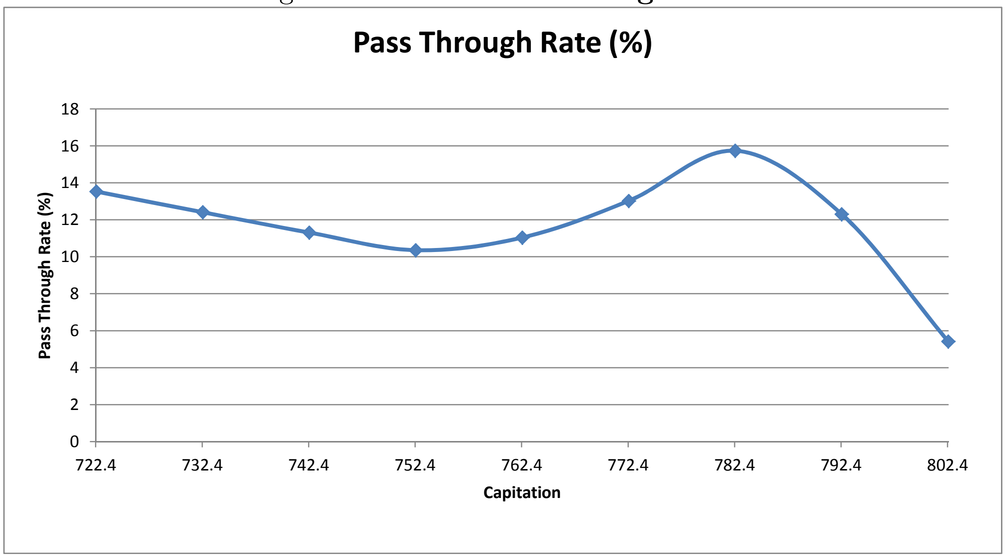
Finally, "product 0" with utility measure $u_0$ in (2.4), represents the option to enroll in the traditional Medicare program as opposed to one of the private MA plans.

Thus far, the out-of-pocket cost for those enrolled in the traditional Medicare program are unaffected by the level or the structure of the capitation rate. The same would apply to several of the reform proposals, for example, the Dominici-Rivlin plan and the so called "Plan One" in Rivlin and Daniel [2015], otherwise advocating a replacement of the exogenous constant capitation rate (R1) by either (R4) or (R5), see (1.8). Under all such programs, the traditional Medicare program is to be treated as the "outside good" or "product 0" with $b_0 = 0$ (PIOG). However, the Wyden-Ryan plan prescribed that the traditional Medicare plan would compete along side with the private MA plans, and this provision is also part of the so-called "Plan Two" option in Rivlin and Daniel [2015]. These types of proposals are to be modelled by setting $b_0 > 0$, more specifically $b_0 = b_1 = \cdots = b_n = b > 0$ (PSOG).

In Federgruen and Lu [2016b], we have applied our model to the Medicare market in 2010, using nation-wide county-by-county data. The model was applied to all 2478 (out of a total of 2727) counties with 2 or more private MA plans. The county-by-county data were used to calibrate the model. We have computed the price and market share equilibria in counterfactual studies, under the above three subsidy schemes in (1.8). On this basis, we have estimated that, if the second lowest bid was adopted in 2010, this would have resulted in a reduction by 16% in federal spending, and an increase by 60$ per month for an average beneficiary enrolled in the traditional Medicare program. (The reduction is considerably larger than when premium and market shares are assumed to remain unaffected by the change of the subsidy scheme.)

As mentioned in the Introduction, we have also used the model to estimate the pass-through rate, associated with an across-the-board reduction of the capitation rate, leaving the remainder of the market structure unchanged, in particular the (PIOG) assumption. Figure 1.1 displays the pass-through rate, on the interval $[722, 802]$. The pass through rates are evaluated by computing the price equilibrium that would arise if the capitation rate was reduced by \$10 from its current

Figure 1.1: **MA Pass-Through Rate**




value. (In all scenarios, the equilibrium was verified to be unique.) Note that the pass through rate fails to be constant, either in absolute or relative terms as implied by a linear or log-linear regression equation, with the premium as the dependent variable and the capitation rate as an explanatory variable. The average pass through rate is 12%, considerably lower than the range of estimates from regression studies, see Section 1.1. In traditional settings, the pass through rate measures what percentage of an increase in the product's cost rate, is passed on to the consumer by an increase in the retail price. In our setting, an increase in the capitation rate, does not affect the profit margin so there is no incentive to increase the retail price. To the extent, modest price increases are implemented, they serve to increase the profit margin, while the consumer's *net price* for the product is *reduced* anyway. However, any firm's unilateral price increase causes many consumers to migrate to other plans, therefore limiting its appeal.

### 1.4.2 Manufacturer Coupons

Manufacturers use coupons to stimulate sales of their products. While paper coupons have been in use for a century, modern delivery methods include internet and so-called mobile coupons. The latter is an electronic ticket delivered to a mobile phone that can be exchanged for a fixed dollar amount rebate, or a percentage discount when purchasing a specific product or service. While newspaper inserts are still the primary method of distribution, the use of internet coupons has mushroomed by 263% in the year 2009 alone, see Wall Street Journal [May 8 2010]. As mentioned in the Introduction, the marketing literature has studied brand choice models, with the relative price as the explanatory variable in the products' utility measures. In other words, customers base their brand choices on the relative price (= nominal price - the coupon value), even though, often, only a minority of customers end up redeeming their coupon. Examples of such brand choice papers with a MNL consumer choice model include Gonul and Smith [1999] and Mela *et al.* [1997].

The model in Section 2.4 applies to these settings, this time with $g_i(\mathbf{p}) = C_i$ or $g_i(\mathbf{p}) = \delta_i p_i$, as in (R1) and (R2), respectively, and with the response function $f(\cdot)$ a general, increasing linear or non-linear function. Product 0, now, represents the no-purchase option, so that $b_0 = 0$, as discussed in Section 2.4.

### 1.4.3 Green Technologies

Many governments are interested in promoting the production and sales of "green" technologies, for example, solar panels and electric vehicles. For the former, sales have been stimulated by offering Feed-In Tariffs, a long term guaranteed purchase agreement for producers to sell their electricity into the grid, see Alizamir *et al.* [2016] and the references therein. As an alternative stimulus mechanism, major subsidies have been provided in the US, both at the federal and state level, mainly in the form of tax rebates. In 2009, for example, the federal government raised the tax credit for solar panels to 30% of the nominal price, as part of the American Recovery and Reinvestment Act(ARRA). Similar programs exist in Europe, in particular in Germany, one of the countries leading the adoption of solar panels, see e.g., Jaeger-Waldau [2007] and Lobel and Perakis [2016].

The ARRA also aimed to stimulate the adoption of electric vehicles. Here, a constant tax rebate of \$7,500 per vehicle was instituted, reducing the effective manufacturer's suggested retail price for, say, the Chevy Volt, from \$39,145 to \$31,615, or the Tesla model S from \$59,350 to \$51,850. See Cohen *et al.* [2016] and King *et al.* [2014] for more detailed descriptions of the green technologies industries.

The model in Section 2.4, can, again, be applied to characterize the competition in the oligopolistic industries of solar panel manufacturers and the electrical automobile industry. The subsidy scheme for solar panels is of type (R2) while that for electric vehicles is of type (R1).

### 1.4.4 Pharmaceuticals

The vast majority of the US population has a medical insurance plan that covers drugs in most categories. Drugs that treat the same medical condition are said to be in the same *category*. Every insurance company specifies within each category, an approved list of drugs, called the *formulary*. Each drug, within the formulary, comes with a co-payment. Even though the co-payment is often a small part of the full cost, it is well known that most consumers are sensitive to differences among them.

As a consequence, many drug manufacturers have offered coupons for specific drugs. Cahn [2012] estimates that the manufacturers spend "between \$3 billion and \$6 billion annually on coupon programs", and that "coupons will be applied to 50 million brand name prescriptions by 2021". As a further indication of the prevalence of the coupon practice, in 2009, coupons were offered for half of the top selling drugs, see Cahn [2012].

The competition among the manufacturers within a given category may be analyzed by the model in Section 2.4. In this case, it is insurance companies that provide a subsidy to the consumers covered by their plans. For a given insurance plan and drug $i$, let

$$\begin{aligned} p_i &= \text{the nominal price, per unit of dosage} \\ copay_i &= \text{the co-payment, per unit of dosage} \\ coupon_i &= \text{the coupon, per unit of dosage} \end{aligned}$$

The manufacturer of product i earns a net reward $= p_i - coupon_i$, while the insurance company offers a subsidy $g_i(\mathbf{p}) = C_i = p_i - copay_i$, see (2.1a), resulting in a net price $= copay_i - coupon_i$ for the consumer. (The net price is sometimes negative.) With nominal or list prices set in advance of the formulary and co-payments, the manufacturers compete by selecting their coupon values $\{coupon_i\}$.

On the global scene, one finds national governments, international organizations and private foundations subsidizing specific drugs and vaccines to combat infectious diseases. Such international organizations include the World Health Organization and UNICEF. The Global Alliance for Vaccines and Immunisation (GAVI) is a public-private global health partnership committed to increasing access to immunisation in poor countries, specifically with the help of subsidy programs. Indeed, GAVI is the primary subsidy source for many vaccines. Finally, the Bill and Melinda Gates Foundation is an example of a private foundation, pursuing the same goals, among other global health campaigns.

Two very specific examples are (i) vaccines to immunize against yellow fever, that has more than 900 million people at risk, and (ii) drugs to combat malaria, a disease afflicting 300-500 million individuals, with 2 million lethal cases, annually. As to the former, there are four global suppliers[2] providing 33 billion vials, annually, while GAVI estimates the global demand to be between 78 and 137 billion vials. To promote the vaccination program in developing countries, GAVI offers a country dependent constant subsidy per vial[3]. Artesiminin Combination Therapies (ACT) represent the most effective drug to treat malaria. In 2007, the World Bank created the Affordable Medicines Facility for malaria (AMFm) to provide a *uniform subsidy* for each treatment unit, again to stimulate the consumption of these drugs. Levi *et al.* [2016] model the competition among the suppliers as a homogenous Cournot competition game with a linear inverse demand function. Subsidies are provided by a central planner, the focus of the paper is when a uniform subsidy per unit is optimal among all, possibly firm-dependent constant subsidies.

### 1.4.5 Prospect Theoretical Models

As mentioned in the Introduction, the price competition game in Section 2.4 also serves as an adequate representation of prospect theoretical models for oligopolies, even when customers pay the full price, without any subsidization. Behavioral economists have demonstrated that consumers evaluate alternative products or services on the basis of relative prices, i.e, their differentials with respect to a given reference point. Moreover, these differentials are weighted in non-linear ways; in particular, it is well established that positive price differentials are weighted more heavily than equal sized negative differentials. To represent this loss aversion phenomenon, often a piecewise

[2]Bio-Manguinhos(Brazil), Institut Pasteur de Dakar (Senegal), FSUE Chumakov (Russia)and Sanofi Pasteur (France).

[3]Except for Latin American countries where subsidies are provided by the Pan American Health Organization(PAHO).

linear, convex response function $f(\cdot)$ of the type:

$$f(x) = \beta x, \text{ if } x < 0 \text{ and } f(x) = \alpha x, \text{ if } x > 0, \text{ with } \beta < \alpha, \tag{1.9}$$

is used. This is the "asymmetric reference price model" adopted in almost all marketing studies, see the Introduction and Mazumdar *et al.* [2005]. In the following sections, we pay special attention to this type of response function among more general nonlinear functions required to represent the phenomenon of *diminished sensitivity*, i.e., the fact that the impact of a \$200 differential may not be double that of a \$100 differential. Our comparison results in Section 1.7 are focused on the structure in (1.9).

## 1.5 Exogenous Constant or Percentage Reference Value

In this section, we study the price competition model with the reference value exogenously set as a constant or a given fraction of the absolute price, i.e., when $g_i(p)$ satisfies (R1) or (R2). We start with the case of constant reference values.

### 1.5.1 Constant Reference Values (R1)

Under constant reference values, the market-shares in (2.5) are given by

$$d_i(\mathbf{p}) = \frac{\exp\left(a_i - b_i f\left(p_i - C\right)\right)}{\exp\left(a_0 - b_0 f\left(p_0 - C\right)\right) + \sum_{k=1}^{N} \exp\left(a_k - b_k f\left(p_k - C\right)\right)}, \quad i = 1, 2, \ldots, N. \tag{1.10}$$

Taking derivatives with respect to the price variables, we get, in view of Assumption 1:

$$\frac{\partial d_i}{\partial p_i} = -b_i f'\left(p_i - C\right) d_i(1 - d_i), \quad \text{if } \Delta p_i = p_i - C \notin \mathbb{P}, \tag{1.11}$$

$$\frac{\partial d_i}{\partial p_j} = b_j f'\left(p_j - C\right) d_i d_j, \quad \text{if } \Delta p_j = p_j - C \notin \mathbb{P} \text{ for any } j = 1, 2, \ldots, N, j \neq i. \tag{1.12}$$

When $\Delta p_i$ or $\Delta p_j \in \mathbb{P}$, formulae, similar to (1.11) and (1.12), apply to the left-derivatives and the right-derivatives of $d_i$ with respect to $p_i$ and $p_j$, respectively. Note that each firm's demand is decreasing in its own price and increasing in any of its competitors' prices, as in the classical MNL model.

**Theorem 1** *Under an exogenously specified subsidy, the competition model is log-supermodular. In particular, there exists a pure Nash equilibrium price vector $p^*$, and the set of all price equilibria is a lattice and, therefore, has a componentwise largest and smallest element, $\bar{p}^*$ and $\underline{p}^*$, respectively.*

An additional implication of the price game being (log-)supermodular is the fact that an equilibrium may be computed with a simple tatônnement scheme: starting with an arbitrary price vector $p^{(0)}$, one iteratively computes a best response price for each of the $N$ firms to the most recently generated prices of the competitors. The scheme is guaranteed to converge to an equilibrium. Moreover, when the tatônnement scheme is started at $p^{min}[p^{max}]$, it is guaranteed to converge to $\underline{p}^*[\bar{p}^*]$. It is therefore possible to unequivocally determine whether the game has a unique equilibrium by starting the tatônnement scheme, both at $p^{min}$ and at $p^{max}$, and checking whether the two schemes converge to the same limit point. We are, at this point, unaware of any, a priori, theoretical conditions which guarantee the uniqueness of the price equilibrium under (R1); see, however, Theorem 3(b). It can also be shown that

**Proposition 1** *Assuming that the response function $f(\cdot)$ is increasing and convex, each firm's profit function is strictly quasi-concave in its own price.*

The quasi-concavity property, of course, greatly simplifies the computation of best-response prices. More specifically, when $f(\cdot)$ is convex, there is a *unique* best response price, for any set of prices selected by the competitors. This best response price, for given choices of firm $i$'s competitors $p^0_{-i}$, can be determined as: $\inf\{p_i : \frac{\partial \log(\pi_i)(p_i, p^0_{-i})}{\partial p_i} < 0\}$.

Next, we show that for $p^{max}$ sufficiently large, any equilibrium $p^*$ is in the *interior* of the feasible price space $\mathbf{X}_{i=1}^N[p_i^{min}, p_i^{max}]$. Note first that

$$\begin{aligned}
\lim_{p_i \searrow c_i} \frac{\partial \log \pi_i(p_i, p_{-i})}{\partial p_i} &= \lim_{p_i \searrow c_i} \left[ \frac{1}{p_i - c_i} + \frac{\partial \log d_i(p_i, p_{-i})}{\partial p_i} \right] \\
&= \lim_{p_i \searrow c_i} \left[ \frac{1}{p_i - c_i} - b_i f'(p_i - C)(1 - d_i) \right] = +\infty,
\end{aligned}$$

since the second term within the squared brackets is bounded in $p_i \searrow c_i$. This implies that, for any equilibrium $p^*$, $p_i^* > c_i$ for all $i = 1, 2, \ldots, N$. Similarly,

$$\begin{aligned}
\lim_{p_i \nearrow \infty} \frac{\partial \log \pi_i(p_i, p_{-i})}{\partial p_i} &= \lim_{p_i \nearrow \infty} \left[ \frac{1}{p_i - c_i} - b_i f'(p_i - C)(1 - d_i) \right] \\
&\leq \lim_{p_i \nearrow \infty} \left[ -b_i f'_+(0)(1 - d_i(p_i, p_{-i})) \right] \\
&\leq -b_i f'_+(0)(1 - d_i(C, p_{-i})) < 0,
\end{aligned}$$

since $d_i$ is decreasing in its own price $p_i$ and $f$ is convex. This implies that, for $p^{max}$ sufficiently large, $p^* < p^{max}$. Thus, if $p^{max}$ is sufficiently large, any equilibrium $p^*$ is an interior point of the feasible price space, so that

$$(p_i^* - c_i)b_i f'\left(p_i^* - C\right)(1 - d_i) = 1, \qquad \text{if } p_i^* - C \notin \mathbb{P}. \tag{1.13a}$$

$$\left.\begin{aligned} (p_i^* - c_i)b_i f'_+\left(p_i^* - C\right)(1 - d_i) \geq 1 \\ (p_i^* - c_i)b_i f'_-\left(p_i^* - C\right)(1 - d_i) \leq 1 \end{aligned}\right\}, \qquad \text{if } p_i^* - C \in \mathbb{P}. \tag{1.13b}$$

This set of equations (or inequalities) may be used, in any empirical study, to infer the firms' cost rates $c$ from the observed price equilibrium in the market. In particular when $p_i^* - C \notin \mathbb{P}$, for all $i = 1, 2, \ldots, N$, the First Order Conditions (1.13) reduce to a system of equations, with the unique solutions:

$$c_i = p_i^* - \frac{1}{b_i f'(p_i^* - C)(1 - d_i(p^*))}, \quad i = 1, 2, \ldots, N. \tag{1.14}$$

When $p_i^* - C \in \mathbb{P}$, for some firm $i$, the cost rate $c_i$ can be determined only within an interval, the width of which depends on the difference between the right hand and left hand derivatives $[f'_+(p_i^* - C) - f'_-(p_i^* - C)]$:

$$p_i^* - \frac{1}{b_i f'_-(p_i^* - C)(1 - d_i(p^*))} \leq c_i \leq p_i^* - \frac{1}{b_i f'_+(p_i^* - C)(1 - d_i(p^*))}. \tag{1.15}$$

We complete this section with an exploration of whether the price equilibrium is *monotone* in the reference value $C$. This conjecture is intuitive: after all, interpreting the reference value as a subsidy to the consumer by a third-party, it appears intuitive that the larger the subsidy, the more firms are incentivized to increase their nominal prices. Assuming the increase in the firms' nominal prices is smaller than the increase in the reference value, i.e., assuming pass through rates of up to 100%, this generates a setting with lower net prices, yet better marginal profit rates. Indeed, the monotonicity property can be proved under the classical diagonal-dominant condition:

$$(D) \qquad \sum_{j=0}^{N} \frac{\partial d_i(p)}{\partial p_j} \leq 0. \quad \forall i = 1, 2, \ldots, n. \tag{1.16}$$

( See (1.11) and (1.12) with the understanding that, when $\frac{\partial d_i(p)}{\partial p_j}$ fails to exist, it is replaced by the left-hand derivative. ) In view of (1.11) and (1.12), the inequality (1.16) is equivalent to

$$|\frac{\partial d_i(p)}{\partial p_i}| \geq \sum_{j \neq i} \frac{\partial d_i(p)}{\partial p_j}. \tag{1.17}$$

i.e., in the matrix $\left(\frac{\partial d_i(p)}{\partial p_j}\right)_{i,j}$, any diagonal element dominates, in absolute value, the sum of its off-diagonal elements. This condition appears intuitive: it states that when all firms in the market reduce their prices by the same amount, this will not result in a *decrease* of any of product's sales volume. The diagonal dominant condition goes back to Arrow *et al.* [1959] and Hadar [1965], and is a standard assumption in many price competition models, see e.g., Vives [2001], Bernstein and Federgruen [2004], Farahat and Perakis [2011] and Allon *et al.* [2013].

**Theorem 2** *Assume that $f$ is convex and (D) applies on the feasible price cube $[p^{min}, p^{max}]$. Then, the component-wise smallest and component-wise largest price equilibrium $\underline{p}^*$ and $\overline{p}^*$ are monotonically increasing in $C$.*

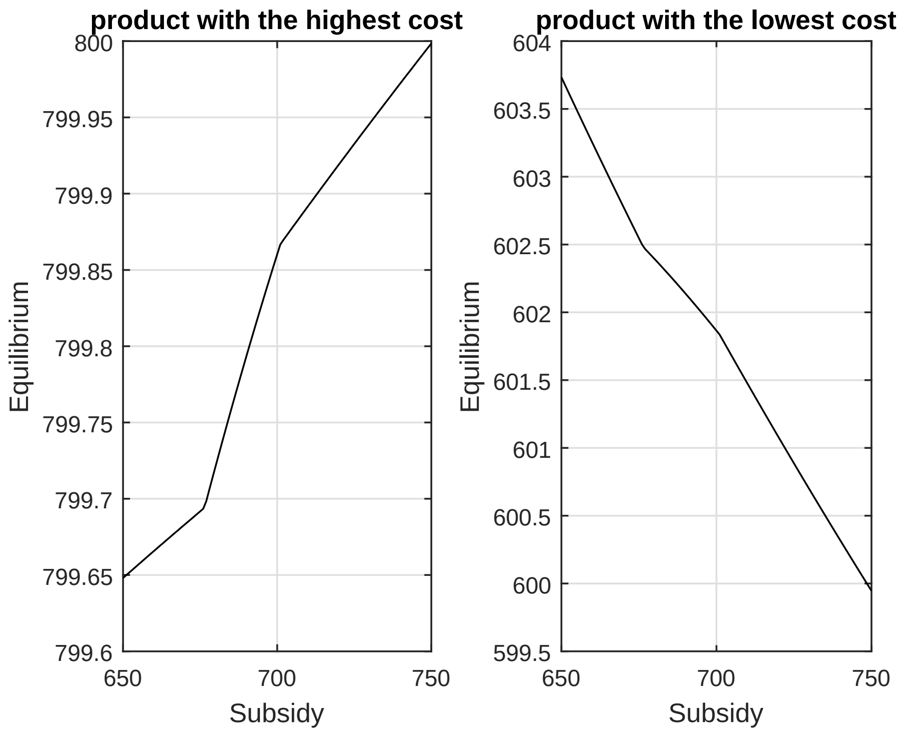
In the standard MNL model, condition (D) is known to hold, see e.g., the proof of Theorem 5 in Bernstein and Federgruen [2004], as well as Proposition 2 below. Under the PIOG-structure, the diagonal dominant condition (D) – and hence monotonicity of the price equilibrium in $C$ – can be shown as well, as long as the market share of the outside good is above a given threshold value, which depends on the *degree of non-linearity* of the response function.

**Proposition 2** (PIOG) *Assume $b_0 = 0, b_1 = \cdots = b_N = b$. If $f$ is convex and $d_0 \geq 1 - \frac{\beta}{\alpha}$ on the feasible price cube $[p^{min}, p^{max}]$, then the diagonal-dominant condition (D) holds, and the component-wise smallest and component-wise largest price equilibrium $\underline{p}^*$ and $\overline{p}^*$ are monotonically increasing in $C$.*

In the standard MNL model, $\alpha = \beta$, so that the market share condition $d_0 \geq 1 - \beta/\alpha = 0$ holds trivially. In our general model, the market share condition easily applies, in many applications. For example, in the above Medicare insurance market, $\alpha = 1, \beta = 0.75$ and the traditional Medicare program, representing the outside good, has a current market share of 70%, well above the 25% $(= 1 - \beta/\alpha)$ threshold required by Proposition 2. The managerial implication in that market is clear: an across-the-board reduction of the capitation rate by a given amount, will result in a decline of all premia. This monotonicity structure is widely assumed by policy makers and underlines the structure of the reduced form regression equations employed in various studies, see the Introduction.

In the PSOG case, where $b_0 = b_1 = \cdots = b_N = b$, the diagonal-dominant condition (D) is harder to guarantee, as the following counterexample demonstrates. The example considers a market with $n = 3$ products, along with an outside good. The model parameters are specified in the caption of Figure 1.2 below. The example shows, that the product with the lowest price adjusts its price downward when the reference value $C$ is increased. ( By Theorem 2, this implies that the diagonal-dominant condition (D) fails in this example. ) Note that, under PSOG, as all net prices decline by the same amount, the outside good gains in attractiveness, along with all in-market products, forcing some products to reduce their normal prices. Indeed, as the reference value increases, the price *spread* increases as well.

Figure 1.2: **Equilibrium Price as a Function of Constant Reference Value**




*Notes:* The parameters are: $N = 3, b_i = 0.0134, i = 0, 1, 2, 3, [a_0, a_1, a_2, a_3] = [0, -3.25, -4.24, -4.70], [c_1, c_2, c_3] = [720.84, 593.16, 487.06], p_0 = 850, f(x) = x^+ - 0.75x^-$. (These parameters are calibrated from Medicare data, see Federgruen and Lu [2016b].) Product 0 is the traditional Medicare program.

### 1.5.2 Percentage Reference Values

Assume, now, that the reference values $\{g_i(\mathbf{p})\}$ are given by (2.1b), for specific fractions $\delta_i < 1$.

**Theorem 3** *(a) Under exogenously specified percentage discount subsidies, the competition model is log-supermodular. This implies that the price game has a pure Nash equilibrium, and that* *the set of equilibria is a lattice, with a componentwise smallest and largest element.*

*(b) If the response function $f(\cdot)$ is convex, the equilibrium is unique.*

In conclusion, under the basic Assumption 1, the price competition model is log-supermodular, both when the reference value is an exogenous constant, and when it is specified as a given percentage of the absolute price. This implies the existence of a component-wise smallest and largest equilibrium, which can be computed with a simple tatônnement scheme. It is difficult to identify sufficient conditions guaranteeing the existence of a unique equilibrium under a constant reference value. In contrast, convexity of the response function provides such a guarantee when the reference value is specified as a discount percentage of the absolute price.

Consider now the case where all reference values are determined as a uniform fraction $\delta$ of the nominal price, that is $\delta p_i$ for all $i$. It is, again, of interest to explore whether a larger reference value due to an increase of $\delta$, results in larger equilibrium prices. We prove this monotonicity result for a response function $f$, which is linear on the positive half line, for all $\delta$ in excess of a minimum threshold value. ( Under (R2), net prices are never negative, rendering the shape of the response function on the negative half line immaterial. )

Let $\overline{\delta} = 1 - \frac{1}{\overline{b}\overline{p} - \underline{b}\underline{p}}$, where $\underline{b} = \min\{b_i, i = 0, 1, 2, \ldots, N\}, \overline{b} = \max\{b_i, i = 0, 1, 2, \ldots, N\}, \underline{p} = \min\{p_i^{min}, i = 0, 1, 2, \ldots, N\}, \overline{p} = \max\{p_i^{max}, i = 0, 1, 2, \ldots, N\}$, the following proposition shows the monotonicity of the price equilibrium with respect to the discount factor $\delta$ when $\delta \geq \overline{\delta}$.

**Proposition 3** *Assume that $f(x) = \alpha x$ for $x \geq 0$ and $\delta_i = \delta$ for all $i$. Then, the unique price equilibrium is monotonically increasing in the discount factor $\delta$ when $\delta \geq \overline{\delta}$.*

The lower bound condition $\delta \geq \overline{\delta}$ is required to provide a theoretical guarantee for the monotonicity result, however, in our extensive numerical studies, we found that the monotonicity prevails at any level of the discount factor $\delta \in [0, 1)$.

## 1.6 The $n$-th Lowest Price Reference Value

In this section, we study the competition model with a reference value endogenously determined as the *lowest* or $n$-th lowest among the nominal prices selected by the competing firms, namely, $g(\mathbf{p}) = p_{(n)}$.

### 1.6.1 Reference Value Determined by The Lowest Price

We start with the case of the lowest price, i.e., reference value structure (R3). Some additional structure is needed for the response function $f(\cdot)$, i.e., we replace Assumption 1 by the a somewhat stronger condition regarding the shape of the response function:

**Assumption 2** *$f(x)$ is increasing, convex and differentiable everywhere with the possible exception of $x = 0$.*

We allow for non-differentiability in $x = 0$, to capture the application of our model to the Medicare market in Section 1.4.1, as well as prospect theoretical models discussed in Section 1.4.5. We first derive expressions for the sales volumes $\{d_i(p) : \ i = 1, 2, \ldots, N\}$ and their derivatives $\{\frac{\partial d_i}{\partial p_i} : \ i = 1, 2, \ldots, N\}$. To this end, we distinguish among (i) the case where $p_i < p_{(1)}^{-i}$, (ii) the case where $p_i > p_{(1)}^{-i}$, and (iii) $p_i = p_{(1)}^{-i}$.

(i) For any $\underline{p_i < p_{(1)}^{-i}}$, the reference value $g(\mathbf{p}) = p_i$. By (2.5) and the fact that $f(0) = 0$, we have

$$d_i(\mathbf{p}) = \frac{\exp(a_i)}{\exp(a_0 - b_0 f(p_0 - p_i)) + \sum_{k=1}^{N} \exp(a_k - b_k f(p_k - p_i))}, \tag{1.18}$$

Taking the derivative with respect to $p_i$ yields

$$\begin{aligned} \frac{\partial d_i}{\partial p_i} &= -\frac{\exp(a_i) \cdot \sum_{k \neq i} \exp(a_k - b_k f(p_k - p_i)) \cdot (b_k f'(p_k - p_i))}{\left(\exp(a_0 - b_0 f(p_0 - p_i)) + \sum_{k=1}^{N} \exp(a_k - b_k f(p_k - p_i))\right)^2} \\ &= -d_i \sum_{k \neq i} d_k b_k f'(p_k - p_i). \end{aligned} \tag{1.19}$$

(ii) For any $\underline{p_i > p_{(1)}^{-i}}$, the reference value $g(\mathbf{p}) = p_{(1)}^{-i}$ and similar to case (i), the market share in (2.5) satisfies

$$d_i(\mathbf{p}) = \frac{\exp\left(a_i - b_i f\left(p_i - p_{(1)}^{-i}\right)\right)}{\exp\left(a_0 - b_0 f\left(p_0 - p_{(1)}^{-i}\right)\right) + \sum_{k=1}^{N} \exp\left(a_k - b_k f\left(p_k - p_{(1)}^{-i}\right)\right)}, \tag{1.20}$$

$$\frac{\partial d_i}{\partial p_i} = -b_i f'\left(p_i - p_{(1)}^{-i}\right) d_i (1 - d_i). \tag{1.21}$$

(iii) When $\underline{p_i = p_{(1)}^{-i}}$, it is easily verified that both (1.18) and (1.20) represent correct expressions for the sales volume $d_i(p)$. Hence, the derivative $\frac{\partial d_i}{\partial p_i}$ may fall to exist for the value $p_i = p_{(1)}^{-i}$. However, the left- and right hand derivatives $\frac{\partial_- d_i}{\partial p_i}$ and $\frac{\partial_+ d_i}{\partial p_i}$ exist; the former is given by (1.19) and the latter by (1.21).

We first demonstrate that when the reference value is determined by the lowest price, the demand volume continues to have the desired monotonicity properties: each product's demand volume decreases in its own price and increases in the price of any of the competing products. While highly intuitive, we will see that these properties may fail under alternative reference value structures, e.g., (R5), see Section 2.7.

**Proposition 4** *Let Assumption 2 hold.*

*(a) Every product's demand volume is a decreasing function of its own price.*

*(b) The products are substitutes, i.e., the demand for any product is increasing in the price of any of the competing products.*

Proposition 4(b) shows that, when the reference value is specified as the lowest price, the products act as substitutes. Note also that if the demand for any product $i$ is non-decreasing in the price of any alternative product, the same monotonicity property applies to the associated profit values, and vice versa. The latter monotonicity property (of the profit functions) is often referred to as "the competitive market" property, see e.g., Assumption 1 in Cabral and Villas-Boas [2005].

In addition to Assumption 2, we assume the following two conditions hold: For any product $i = 1, 2, \ldots, N$, and any price vector $p_{-i}$,

$$\sum_{j \neq i} \frac{\partial_- d_i}{\partial p_j}(p_{(1)}^{-i}, p_{-i}) \leq |\frac{\partial_+ d_i}{\partial p_i}(p_{(1)}^{-i}, p_{-i})|, \qquad (\tilde{D})$$

$$(p_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial p_j} \text{ is quasi-convex in } p_i \in [c_i, p_{(1)}^{-i}). \qquad (M)$$

Condition ($\tilde{D}$) is a weaker version of condition (D), a classical dominant-diagonal condition, employed in Section 1.5. Recall that it merely precludes that a uniform price increase by all firms would result in an increase of any of the firms' sales volumes. Moreover, under $\tilde{D}$, the dominant diagonal condition is required only when the firm's price is close to the lowest price offered by the competitors. In Lemma 2 below, we show that conditions ($\tilde{D}$) and (M) apply when $f$ is a two-part piece-wise linear function and all products share the same price sensitivity coefficient (PSOG).

We now show that a pure-strategy price equilibrium exists in the lowest subsidy model by showing that each firm's profit function is quasi-concave in its own price .

**Theorem 4** *Assume that Assumption 2 and conditions ($\tilde{D}$)-(M) apply in the lowest subsidy model. A pure-strategy Nash equilibrium exists.*

The following lemma shows that conditions ($\tilde{D}$) and (M) apply when the function $f(\cdot)$ is a two-part piece-wise linear function and the price sensitivity coefficient is the same for all firms.

**Lemma 2** *The conditions of Theorem 4, i.e., conditions ($\tilde{D}$)-(M) and Assumption 2 apply, when (i) $b_i = b$ for all $i$ and (ii) $f$ is a two-part piece-wise linear function, i.e., $f(x) = \alpha x^+ - \beta x^-$ with $\alpha, \beta > 0$, see* (1.9).

We now show that under the conditions of Lemma 2, the model is (log-)supermodular.

**Theorem 5** *Assume (1) $b_i = b$ for all $i$ and (2) $f(x) = \alpha x^+ - \beta x^-$ with $\alpha, \beta > 0$. The price competition game with the reference value specified as the lowest price is log-supermodular.*

Following the arguments provided in the previous section, it is easily verified that any price equilibrium must be an interior point of the feasible price region, provided the upper bounds $p^{max}$ are sufficiently large. Moreover, as shown in (A-11), each profit function $\pi_i(p_i, p_{-i})$ is differentiable everywhere, with the possible exception of the point $p_i = p_{(1)}^{-i}$. Together with (A-12) and (A-13), this implies that any equilibrium $p^*$ satisfies the following system of equations and inequalities:

$$1 - (p_i^* - c_i) \sum_{k \neq i} d_k b_k f'\left(p_k^* - p_i^*\right) = 0, \quad \text{if } p_i^* < p_{(1)}^{-i}, \tag{1.22a}$$

$$1 - (p_i^* - c_i) b_i f'\left(p_i^* - p_{(1)}^{-i}\right)(1 - d_i) = 0, \quad \text{if } p_i^* > p_{(1)}^{-i}\ , \tag{1.22b}$$

$$\frac{\partial_+ \log \pi_i(p_{(1)}^{-i}, p_{-i}^*)}{\partial p_i} = 1 - (p_{(1)}^{-i} - c_i) b_i f'_+(0)(1 - d_i) \leq 0, \tag{1.22c}$$

$$\frac{\partial_- \log \pi_i(p_{(1)}^{-i}, p_{-i}^*)}{\partial p_i} = 1 - (p_{(1)}^{-i} - c_i) \sum_{k \neq i} d_k b_k f'_+\left(p_k^* - p_{(1)}^{-i}\right) \geq 0. \tag{1.22d}$$

Similar to the system of equations and inequalities in (1.13), (1.22) allows us to determine the cost rates $\{c_i\}$ from any observed price equilibrium $p^*$. For any product with $p_i^* \neq p_{(1)}^{-i}$, the unique corresponding marginal cost rate $c_i$ could be derived from equation (1.22a) or (1.22b). If $p_i^* = p_{(1)}^{-i}$, an interval can be determined for the corresponding marginal cost rate $c_i$ by inequalities (1.22c) and (1.22d).

### 1.6.2 Reference Value Determined by the $n$-th Lowest Price

In this subsection, we study the price competition model when the reference value is endogenously determined as the $n$th-lowest price among the set of nominal prices selected by the competing firms, i.e., $g(\mathbf{p}) = p_{(n)}$, for some $n \geq 2$, see (R4). The Medicare market in Subsection 1.4.1 provides the motivation for this generalization of the lowest price subsidy; as mentioned, several bipartisan proposals, in particular the Wyden–Ryan and Domenici–Rivlin's plans, advocate setting the capitation rate as the *second* lowest price. As above, we assume that Assumption 2 applies.

Note that the reference value satisfies the following relationships:

$$g(p_i, p_{-i}) = p_{(n)} = \begin{cases} p^{-i}_{(n-1)}, & p_i \leq p^{-i}_{(n-1)} \\ p_i, & p^{-i}_{(n-1)} < p_i \leq p^{-i}_{(n)} \\ p^{-i}_{(n)}, & p_i > p^{-i}_{(n)} \end{cases} \tag{1.23}$$

Substituting (1.23) into (2.5), we get the following expressions for the sales volumes:

$$d_i(p_i, p_{-i}) = \begin{cases} \frac{\exp\left(a_i - b_i f\left(p_i - p^{-i}_{(n-1)}\right)\right)}{\sum_{k=0}^{N} \exp\left(a_k - b_k f\left(p_k - p^{-i}_{(n-1)}\right)\right)}, & p_i \leq p^{-i}_{(n-1)} \\ \frac{\exp(a_i)}{\sum_{k=0}^{N} \exp(a_k - b_k f(p_k - p_i))}, & p_i \in (p^{-i}_{(n-1)}, p^{-i}_{(n)}] \\ \frac{\exp\left(a_i - b_i f\left(p_i - p^{-i}_{(n)}\right)\right)}{\sum_{k=0}^{N} \exp\left(a_k - b_k f\left(p_k - p^{-i}_{(n)}\right)\right)}, & p_i > p^{-i}_{(n)} \end{cases}, \tag{1.24}$$

As in the previous section, we need the following variant of condition (D) and (M) which we now refer to as condition (D′) and (M′)

$$|\frac{\partial_+ d_i}{\partial p_i}(p^{-i}_{(n)}, p_{-i})| \geq \sum_{j \neq i} \frac{\partial_- d_i}{\partial p_j}(p^{-i}_{(n)}, p_{-i}), \qquad (D')$$

$$\sum_{j \neq i} \frac{\partial_+ d_i}{\partial p_j}(p^{-i}_{(n-1)}, p_{-i}) \geq |\frac{\partial_- d_i}{\partial p_i}(p^{-i}_{(n-1)}, p_{-i})|,$$

$$(p_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial p_j} \text{ is non-decreasing in } p_i \in (p^{-i}_{(n-1)}, p^{-i}_{(n)}), \qquad (M')$$

The first inequality in $(D')$ is identical to $(\tilde{D})$. Lemma 3, below, provides sufficient conditions for both inequalities in $(D')$.

Similar to Theorem 4, we show that a pure-strategy price equilibrium exists by showing that each firm's profit function is quasi-concave in its own price for any given price choices of the other alternatives.

**Theorem 6** *Assume that Assumption 2 and conditions* $(D') - (M')$ *hold. A pure-strategy price equilibrium exists, when the subsidy is determined as the nth-lowest price.*

The following Lemma shows that conditions $(D') - (M')$ hold under the same conditions as in Lemma 2, for $(\tilde{D})$ and $(M)$, plus an additional assumption which stipulates that a positive relative price weighs more heavily than a negative relative price of the same magnitude. (In prospect theoretical models, the assumption reflects the fact that losses weigh more heavily than equal sized gains.) The assumption $\alpha \geq \beta$ applies in the Medicare insurance market, where each beneficiary pays the full excess of the premium above the subsidy, but receives only part of any shortfall, when the premium is lower than the subsidy.

**Lemma 3** *Conditions* $(D') - (M')$ *hold when (i)* $b_i = b$ *for all* $i = 1, \ldots, N$ *and (ii)* $f$ *is a two-part piece-wise linear function, i.e.,* $f(x) = \alpha x^+ - \beta x^-$ *with* $\alpha \geq \beta \geq 0$.

We conclude that the price competition model has a pure-strategy Nash equilibrium, under minor technical conditions, irrespective of whether the reference value is determined as the lowest price or the $n$-th lowest price. The question remains whether the same fundamental invariance with respect to the subsidy structure applies to the stronger property of the price competition game being (log-)supermodular. It follows from the proofs of Theorem 5 that this (log-)supermodularity property applies whenever each firm's sales volume increases with its alternatives' prices, that is, when products may be viewed as simple *substitutes*, or, equivalently, markets maybe viewed as competitive, see Assumption 1 in Cabral and Villas-Boas [2005]. In the absence of any subsidy, the MNL model clearly represents products that are strict substitutes. Theorem 5 shows that the same applies in our model, when the reference value is determined as the lowest price, subject to minor technical conditions, see Lemma 2(b). However, under a more complex subsidy structure, such as one based on the *second* lowest price, products may cease to interact as strict substitutes; in particular, an increase of the second lowest price in the market, now, has two opposite effects: on the one hand, the *net* prices of all other products decrease by the same amount, by itself resulting in an increase of each of their market shares. However, the non-linearity of the function $f$, even in its simplest form when $f(x) = \alpha x^+ - \beta x^-$, implies a smaller increase of the utility attributed to the cheapest product as supposed to the utility increases for the other alternatives. This has the opposite effect of shifting some of the market share from the lowest priced product toward those other alternatives; the net effect of an increase of the second lowest price, may therefore involve a *decrease* of the market share of the cheapest alternative. Indeed, this phenomenon may well occur depending on the response function's degree of non-linearity $\beta/\alpha$ and the relative market shares of the cheapest and the second cheapest products. More specifically, one can show

$$\frac{\partial d_{(1)}(\mathbf{p})}{\partial p_{(2)}} \geq 0 \quad \text{iff} \quad d_{(2)}(\mathbf{p}) \geq \left(1 - \frac{\beta}{\alpha}\right)\left(1 - d_{(1)}(\mathbf{p})\right),$$

where $d_{(n)}(p)$ denotes the sales volume of the (lowest indexed) product with the $n$-th smallest price.

## 1.7 Comparison of Subsidy Schemes

In this Section, we derive several comparison results across different reference value structures. These results have important implications for the design of subsidy schemes, among other applications. We focus on the comparisons of the price equilibria between reference value structures (R1) and (R2), and (R1) and (R3), in Subsection 1.7.1 and 1.7.2, respectively. To simplify and unify our results, we confine ourselves, in this section, to the asymmetric reference price model, i.e., a two-part piecewise linear response functions and settings where all in-market products share the same price sensitivity coefficient. [(C1) $b_1 = b_2 = \cdots b_N = b$ and (C2) $f(x) = \alpha x^+ - \beta x^-$ with $\alpha \geq \beta > 0$.] Recall that under these conditions, the price competition is log-supermodular under any one of the reference value structures (R1)–(R3).

### 1.7.1 R1 vs R2: Constant vs Discount Based Reference Value

In this subsection, we compare the price equilibrium that arises when the reference value is an exogenous constant, i.e., $g_i(\mathbf{p}) = C$, $i = 0, 1, 2, \ldots, N$, as in (R1), with reference values, specified as given fractions of the absolute prices, i.e., $g_i(\mathbf{p}) = \delta_i p_i$, $i = 0, 1, 2, \ldots, N$, as in (R2).

Let

$$\underline{\delta_i} = 1 - \frac{\beta}{\alpha} \cdot \frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i + b\beta(C - c_i)}}$$

Obviously, $\underline{\delta_i} \in (1 - \frac{\beta}{\alpha},\ 1)$.

**Theorem 7** *Fix $C > 0$; assume $\delta_i \geq \underline{\delta_i}$ for all $i = 1, 2, \ldots, N$.*

*(a) For any given price vector of the competing products, each firm's best response under the discounted price reference value scheme is larger than its best response under the exogenous constant reference value:*

$$p_i^{*DISC}(p_{-i}) \geq p_i^{*EXO}(p_{-i}), \quad \text{for any firm } i \text{ and any } p_{-i}. \tag{1.25}$$

*(b) The unique price equilibrium ($p^{*DISC}$) under the discounted reference value scheme is componentwise larger than any price equilibrium ($p^{EXO}$) under the exogenous constant reference value $C$.*

**Proof.** (a) We show that, for any given price vector $\mathbf{p}$, the best response price vector under the discounted reference value scheme is larger than the best response under the constant reference value. A unique best response vector exists under both schemes, since each firm $i$'s profit function $\pi_i$ is strictly quasi-concave in its own price $p_i$, under both reference value schemes. Proposition 1 shows this property for scheme (R1), while Proposition 2 in Gallego *et al.* [2006] establishes the property for scheme (R2).

Under the discounted reference value scheme, the market share for firm $i$ is given by

$$d_i^{DISC}(\mathbf{p}) = \frac{\exp(a_i - b\alpha\gamma_i p_i)}{\exp(a_0 - b_0\alpha\gamma_0 p_0) + \sum_{j=1}^N \exp(a_j - b\alpha\gamma_j p_j)},$$

where $\gamma_i = 1 - \delta_i$. Under the constant reference value scheme, the market share for firm $i$ is given by

$$d_i^{EXO}(\mathbf{p}) = \frac{\exp(a_i - bf(p_i - C))}{\exp(a_0 - b_0 f(p_0 - C)) + \sum_{k:p_k \geq C} e^{a_k - b\alpha(p_k - C)} + \sum_{k:p_k < C} e^{a_k - b\beta(p_k - C)}},$$

where $f(p_i - C) = \alpha(p_i - C)$ if $p_i \geq C$ and $f(p_i - C) = \beta(p_i - C)$ if $p_i < C$. Note that,

$$\frac{\partial \log \pi_i^{DISC}(p_i, p_{-i})}{\partial p_i} = \frac{1}{p_i - c_i} - b\alpha\gamma_i\left(1 - d_i^{DISC}(p_i, p_{-i})\right)$$

$$\frac{\partial_+ \log \pi_i^{EXO}(p_i, p_{-i})}{\partial p_i} = \begin{cases} \frac{1}{p_i - c_i} - b\alpha\left(1 - d_i^{EXO}(p_i, p_{-i})\right), & p_i \geq C \\ \frac{1}{p_i - c_i} - b\beta\left(1 - d_i^{EXO}(p_i, p_{-i})\right), & p_i < C \end{cases}$$

Given the log-concavity of both functions $\pi_i^{DISC}(p_i, p_{-i})$ and $\pi_i^{EXO}(p_i, p_{-i})$ in $p_i$, for any given vector $p_{-i}$, it suffices to show that

$$\begin{aligned} 0 &\leq \frac{\partial \log \pi_i^{DISC}(p_i, p_{-i})}{\partial p_i} - \frac{\partial_+ \log \pi_i^{EXO}(p_i, p_{-i})}{\partial p_i} \\ &= b\alpha \cdot \begin{cases} \left(1 - d_i^{EXO}(p_i, p_{-i})\right) - \gamma_i\left(1 - d_i^{DISC}(p_i, p_{-i})\right), & p_i \geq C \\ \frac{\beta}{\alpha}\left(1 - d_i^{EXO}(p_i, p_{-i})\right) - \gamma_i\left(1 - d_i^{DISC}(p_i, p_{-i})\right), & p_i < C \end{cases} \end{aligned} \tag{1.26}$$

The proof of (1.26) is relegated to Appendix.

(b) Let $\Psi^{EXO}: \mathbb{R}^N \longrightarrow \mathbb{R}^N$ denote the best response operator in the competition game with an exogenous constant reference value $C$, i.e., $\Psi_i^{EXO}(p)$ is defined as the unique best response for firm $i$ to the price vector $p_{-i}$ in the game. Similarly, let $\Psi^{DISC}$ denote the best response operator in the competition game when the reference value is specified as in (R2). (Best responses are uniquely determined, as shown in part (a).) For any $r \geq 1$, let $\Psi^{EXO(r)}\left[\Psi^{DISC(r)}\right]$ denote the $r$–fold application of the best response operator $\Psi^{EXO}\left[\Psi^{DISC}\right]$, i.e., $\Psi^{EXO(r)}(p) = \Psi^{EXO}\left(\Psi^{EXO(r-1)}(p)\right)$ $\left[\Psi^{DISC(r)}(p) = \Psi^{DISC}\left(\Psi^{DISC(r-1)}(p)\right)\right]$.

Since $f$ is convex, the equilibrium under (R2) is unique, see Theorem 3. It suffices to prove the result for part (b) of the component-wise largest equilibrium $\bar{p}^{*EXO}$. Note that,

$$p^{*DISC} = \lim_{r\nearrow\infty} \Psi^{DISC(r)}(p^{max}) \leq \lim_{r\nearrow\infty} \Psi^{EXO(r)}(p^{max}) = \bar{p}^{*EXO}.$$

The two equalities follow from Theorem 4.3.2 in Topkis [1998], since the price competition game is (log-)supermodular under both reference value schemes, by Theorems 1 and 3. To prove the inequality, we show, by induction that

$$\Psi^{DISC(r)}(p^{max}) \leq \Psi^{EXO(r)}(p^{max}) \quad \text{for any } r = 1, 2, \ldots.$$

For $r = 1$ the inequality holds by part (a) of this theorem. Assume the inequality holds for some integer $r \geq 1$. Then,

$$\begin{aligned}
\Psi^{DISC(r+1)}(p^{max}) &= \Psi^{DISC}\left(\Psi^{DISC(r)}(p^{max})\right) \\
&\leq \Psi^{DISC}\left(\Psi^{EXO(r)}(p^{max})\right) \\
&\leq \Psi^{EXO}\left(\Psi^{EXO(r)}(p^{max})\right) \\
&= \Psi^{EXO(r+1)}(p^{max}).
\end{aligned}$$

The first inequality follows from the fact that the best response operator in a (log-)supermodular game is a monotone operator, while the second inequality follows from part (a) of the theorem. □

**Remark 1** *Theorem 7 holds for both PIOG (i.e., $b_0 = 0$ ) and PSOG (i.e., $b_0 > 0$). In the PIOG case, the minimum discount factor $\underline{\delta_i}$ is increasing in $C$. In the PSOG case, the monotonicity property fails: $\underline{\delta_i}$ is increasing in $C$ if $C \geq p_0$, and decreasing in $C$ when $C < p_0$.*

Theorem 7 leaves us with the question, how the equilibria $p^{*EXO}$ and $p^{*DISC}$ compare, when (some of) the discount percentages $\delta_i$ fall below their threshold value $\underline{\delta}_i$. We focus on the case where all $\delta_i = \delta, i = 1, 2, \ldots, N$. In our numerical results, we have observed that, for any given firm $i$, the following threshold result applies: there exists a threshold value $\overline{\delta^*}_i$ such that

$$p_i^{*DISC} \leq \overline{p_i}^{*EXO} \qquad \text{iff } \delta \leq \overline{\delta^*}_i. \tag{1.27}$$

A similar result prevails for the componentwise smallest equilibrium $\underline{p_i}^{*EXO}$. These results follow from Theorem 7 and the monotonicity of $p^{*DISC}$ in $\delta$, a property we proved for $\delta$ sufficiently large, see Proposition 3, and numerically observe to hold on the entire interval $\delta \in [0, 1)$. Note that when $\delta \searrow 0$, $p^{*DISC}$ approaches the (unique) equilibrium in the absence of subsidization, which also corresponds with $p^{*EXO}$ when the exogenous reference value equals zero. Thus, in general, for sufficiently small $\delta$, $p^{*DISC}$ falls below $\overline{p}^{*EXO}$, see Theorem 2 and Proposition 2. The threshold result (1.27) now follows immediately from the monotonicity of $p^{*DISC}$ in $\delta$ and Theorem 7.

We have characterized the comparison result of the price equilibria under (R1) and (R2), as a function of $\delta$, for a fixed value of the constant reference value $C$. Conversely, it is of interest to characterize the relation for a fixed discount percentage $\delta$, as the constant reference value $C$ varies. Here, we prove a rigorous threshold result, for the most common (PIOG) structure.

**Corollary 1** *Assume (PIOG) and $d_0 \geq 1 - \frac{\beta}{\alpha}$ on the feasible price space. Fix $\delta_i \in [0, 1]$. For each $i = 1, 2, \ldots, N$, there exists a threshold value $C_i^0$ such that $\overline{p_i}^{*EXO} \leq p_i^{*DISC}$ iff $C \leq C_i^0$. A similar threshold result holds for the componentwise smallest equilibrium price $\underline{p_i}^{*EXO}$.*

### 1.7.2 R1 vs R3: Constant vs Lowest Price Reference Value

In this subsection, we compare the price equilibria under an exogenous constant reference value (R1) with those that arise when the reference value is determined as the lowest (nominal) price, i.e., (R3). Here, it is, again, useful to provide a separate treatment for the (PIOG) and the (PSOG) cases. We start with (PIOG), i.e., the most common structure.

Let $\bar{p}^{*EXO}$ ($\underline{p}^{*EXO}$) and $\bar{p}^{*LOW}$ ($\underline{p}^{*LOW}$) be the componentwise largest (smallest) price equilibrium under (R1) and (R3), respectively. We now show that, for any product, the largest equilibrium price under the minimum price reference value is componentwise smaller than the largest equilibrium price under an exogenous constant reference value, if and only if the latter is larger than a given threshold value.

**Theorem 8** (PIOG) *Assume $b_0 = 0$ and $b_1 = b_2 = \cdots = b_N = b > 0$:*

*(a) If $C \leq c_{min}$, $\bar{p}^{*LOW} \geq \bar{p}^{*EXO}$ and $\underline{p}^{*LOW} \geq \underline{p}^{*EXO}$.*

*(b) Assume $d_0 \geq 1 - \beta/\alpha$ on the feasible price region. Fix $i = 1, 2, \ldots, N$, there exists a critical threshold level $C_i^0 \leq \infty$ for the exogenous reference value such that*

$$\bar{p_i}^{*EXO} \geq \bar{p_i}^{*LOW} \quad \text{iff } C \geq C_i^0.$$

*The same threshold result applies to the componentwise smallest equilibria.*

Thus, in the (PIOG) structure, the price equilibria are lower under (R3), as compared to (R1), only if the exogenous constant reference value $C$ is in excess of a given threshold value. Under (PSOG), this dominance relationship prevails, throughout, even when the constant exogenous reference value is relatively low. In Theorem 9, below, we prove the dominance relationship, specifically for a constant reference value $C$ below a given threshold; this result can be extended for larger, hence, arbitrary values of $C$, as long as the component-wise smallest and largest equilibria under (R1) is known to be increasing in $C$. This is guaranteed to be the case under the dominant-diagonal condition (D), see Theorem 2. Recall, however, from the discussion and counter-example in Section 1.5, that monotonicity of the price equilibrium in the constant reference value may fail to hold.

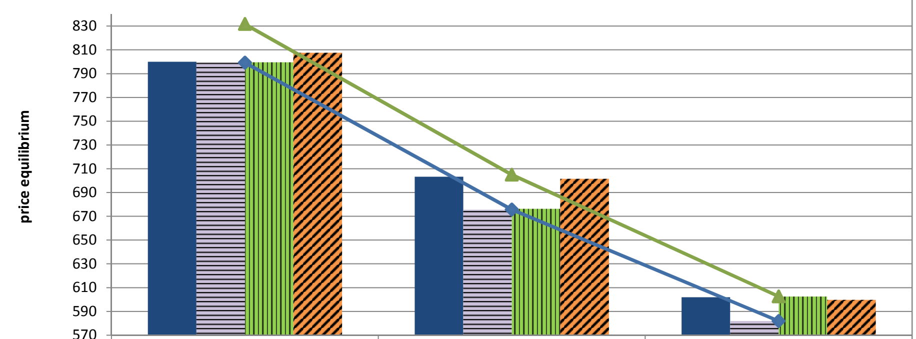
**Theorem 9** (PSOG) *Assume that $b_i = b, i = 0, 1, 2, \ldots, N$.*

*(a) $\underline{p}^{*LOW} \leq \underline{p}^{*EXO}$ and $\overline{p}^{*LOW} \leq \overline{p}^{*EXO}$ if $C \leq c_{min} + \frac{\ln(\alpha)-\ln(\beta)}{b(\alpha-\beta)}$.*

*(b) Assume the diagonal-dominant condition (D) holds. Then, $\underline{p}^{*LOW} \leq \underline{p}^{*EXO}$ and $\overline{p}^{*LOW} \leq \overline{p}^{*EXO}$ for any constant reference value $C$.*

We illustrate the above comparison results with the help of the example used in Figure 1.2. (Recall, this example reflects a market with three products, in addition to the "outside" good.) Note that, the discounted reference value with discount fraction $\delta_i = 0$ for all $i$ results in the well-known standard MNL model. As proven in Theorem 8, the equilibrium prices under a constant reference value $C = \$750$ are componentwise higher than those under a reference value given by the lowest price. However, this ranking breaks down when the reference value is given by the second–lowest price. (Product-3's price under the constant reference value is smaller than when the reference value is given by the second–lowest price). The equilibrium prices under a percentage discount reference value are componentwise smaller than those under a constant reference value when this discount percentage is sufficiently small. This corresponds with the conditions in Theorem 7.

Figure 1.3: **Comparison of Equilibrium Prices under Different Reference Value Structures**




| | product-1 | product-2 | product-3 |
|---|---|---|---|
| Constant | 800.02 | 703.24 | 601.94 |
| Lowest | 799.32 | 675.68 | 581.90 |
| 2nd-Lowest | 799.70 | 676.43 | 602.54 |
| Weighted Average | 807.62 | 701.73 | 599.73 |
| Discount 0% | 799.32 | 675.68 | 581.90 |
| Discount 30% | 831.72 | 705.02 | 602.66 |



*Notes:* The parameters are the same as in Figure 1.2, in addition, the constant reference value $C = 750$, the weighted average reference value structure uses weights $(w_0, w_1, w_2, w_3) = (75.51\%, 5.72\%, 6.86\%, 11.90\%)$. These weights reflect the prior year's market shares of the Medicare plans in this county.

## 1.8 Conclusions and Extensions

In this paper, we have analyzed a general price competition model for settings where the (random) utility attributed to a given product depends on its price, via a non-linear response function $f(\cdot)$ of the *net* price, i.e., the *differential* between the nominal price and a reference value $g_i(p)$, see equations (2.2) and (2.4). Our model is motivated both by settings where the product or service is subsidized by a third party and $g_i(p)$ represents the subsidy level, as well as prospect theoretical models, in which consumers react to *relative* prices, as opposed to absolute prices. In Section 1.4, we have discussed five different application areas for these models.

We have characterized the equilibrium behavior of the price competition model under various specifications of the reference value $g_i(p)$, addressing each of questions (I)-(IV) in the Introduction. In particular, we have derived comparison results across different reference value structures, and different parameters within the same reference structure. These comparison results provide important insights, for the design of subsidy schemes, see e.g. Section 1.4.1, addressing the current debate about the Medicare Advantage Market.

Another common reference value structure is (R5), i.e., the reference value is a weighted average of the selected absolute prices, see for example, Subsection 1.4.1 ad 1.4.5 addressing the Medicare Advantage market and prospect theoretical models, respectively. As mentioned in the Introduction, under this structure, the products (may) cease to be substitutes in the sense of cross price elasticities being positive. Substituting $g_i(\mathbf{p}) = \sum_k w_k p_k$ (, see (2.1e), ) into (2.5) and taking the partial derivative with respect to $p_j$, we obtain:

$$\frac{\partial d_i}{\partial p_j} = d_i \left[ w_j g_i + (1 - w_j) d_j g_j - w_j \sum_{k \neq j} d_k g_k \right], \quad i \neq j, \tag{1.28}$$

where $g_k = b_k f'(p_k - \sum_\ell w_\ell p_\ell), k = 0, 1, \ldots, N$. Thus, for any $i \neq j$,

$$\frac{\partial d_i}{\partial p_j} < 0 \Longleftrightarrow w_j g_i + (1 - w_j) d_j g_j - w_j \sum_{k \neq j} d_k g_k < 0 \Longleftrightarrow w_j (g_i - \bar{g}) + d_j g_j < 0, \tag{1.29}$$

where $\bar{g} = \sum_{k=0}^N d_k g_k \in (\min\{g_k\},\ \max\{g_k\})$ is weighted average of the quantities $\{g_k\}$. In other words, even with identical price sensitivity coefficients $b_1 = b_2 = \cdots = b_N$, a *negative* cross price elasticity may easily arise; in particular, an increase in the price of a product $j$, with a modest market share (i.e., relatively modest $d_j$ value), and hence an increase in the reference value $g(\mathbf{p})$, may end up *decreasing* the market share of a low–cost competing product $i$ (, with a relatively low value of $g_i$.). This phenomenon arises, in particular, under significantly non-linear and convex response function $f(\cdot)$. Alternatively, even when the response function $f(\cdot)$ is linear, negative cross elasticities may arise under heterogenous price sensitivity coefficients $\{b_k\}$.

As mentioned in the Introduction, when products fail to be substitutes throughout the full price region, the existence of a Nash equilibrium can, in general, not be guaranteed. Indeed, elementary structural properties of the profit functions, such as (log-)supermodularity or (log-)quasi-concavitity fail to hold. In terms of the latter, one easily verifies that:

$$\frac{\partial \log(\pi_i)}{\partial p_i} = \frac{1}{p_i - c_i} - \left[ (1 - w_i)(1 - d_i)g_i + w_i \sum_{k \neq i} d_k g_k \right].$$

This function fails in general to be decreasing, i.e., the profit function $\pi_i$ fails to be log-concave unless the products are substitutes (i.e., $\frac{\partial d_k}{\partial p_i} > 0$ for $k \neq i$) and the quantities $\{g_k\}$ are independent of the price $p_i$. As discussed above, this situation arises only when the response function $f(\cdot)$ is linear and the price sensitivity coefficients $\{b_k\}$ are identical.

Future work should address various generalizations of the MNL–based consumer choice model. First, the specification of the utility measure in (2.2) and (2.4) is broad and captures almost all applications we are aware of. However, in some studies a variant of the specification is desired, thus requiring an adaptation of the analysis in the paper. For example, in some applications, one may wish to specify the utility measure as a function of *both* the absolute and the relative price, for example, $u_{i,j} = a_i - \eta_i p_i - b_i f(p_i - g_i(p)) + \varepsilon_{i,j}$, see Chen *et al.* [2016] for such a specification.

Other generalizations relate to Mixed MultiNomial Logit models and nested MNL models. The former are important in applications where the market needs to be partitioned into segments, each with its own utility measures $\{u_i(\mathbf{p})\}$. For example, econometric models for the Medicare Advantage market, e.g., Nosal [2012], identify inertia and switching costs, as a major factor in describing the industry dynamics. To model this phenomenon, the market needs to be segmented into $N$ segments, based on the beneficiary's plan choice in the prior year. The same applies to prospect theoretical models in which the reference value is given by the price of the most recently purchased brand. In addition, socio-economic factors may induce the need for segmentation. Generalizations of our results to *nested* MNL model are required to understand the equilibrium behavior in one of the models used by Curto *et al.* [2015].

# Chapter 2

# Medicare Reform: Estimation of the Impacts of Premium Support Systems

Awi Federgruen and Lijian Lu

## 2.1 Introduction and Summary

Medicare provides health insurance to all US citizens and permanent residents, ages 65 and older, as well as younger people with specific disabilities. The current Medicare system was put in place in 2003, with the adoption of the Medicare Modernization Act (MMA), and the adoption of Medicare Advantage (MA), formerly known as Medicare Choice or Medicare Part C. In 2015, Medicare has covered more than 55 million individuals, at an annual cost of approximately 600 billion dollars. Moreover, without any restructuring, Medicare costs are estimated to grow at twice the rate of the GDP, the result of the upcoming retirement of many baby boomers, increased longevity, as well as the escalating costs of healthcare. The Congressional Budget Office has estimated that the government's aggregate healthcare liabilities, as a percentage of the GDP, would grow from 5% to 12% in the next 40 years, in the absence of a fundamental restructuring of the system. It is generally understood that this would bankrupt the Medicare system.

Prominent members of Congress, bipartisan policy centers and think tanks have launched proposals to rescue the program. A major part of these proposals consists of changes in the way capitation rates or subsidies for the insurance premia would be determined. Some advocate reducing the *exogenously* pre-specified capitation rates, while others call for an *endogenous* determination as a function of the premium bids of all plans competing in the county, for example the *lowest* bid, the *second lowest* bid or a *weighted average* of the bids. Attempts to quantify the implications of such reforms have been hampered by the fact that the equilibrium behavior of the resulting price competition games has not been known. As a consequence, estimates of cost savings for the Medicare program or out-of-pocket costs for the beneficiaries, among other performance measures, have been made, on the basis of the assumption that all plans premium bids and, hence, all market shares would not be affected by the modified subsidy schemes. See e.g. Congressional Budget Office (2013).

Even the impacts of mere *level* changes within the current system with exogenous capitation rate are difficult to predict. Song *et al.* [2012b], Song *et al.* [2013], Cabral *et al.* [2014], Duggan *et al.* [2014], and Curto *et al.* [2015] have applied regression models to estimate the pass-through rates, i.e., the absolute change in the premia due to changes in the capitation rates, with estimated pass through rates varying between 37% and 100% across the different studies, leaving policy makers with a great deal of ambiguity. However, in Federgruen and Lu [2016a] we have *calculated* the pass through rates that result from equilibrium charges within fitted competition game models. Our results show: (1) the pass through rate fails to be constant, neither in absolute nor in relative terms, as implied, respectively, by a linear or log-linear regression model. Instead, as the capitation rate is reduced by up to $30 from their prevailing average ($802), the pass through rate varies in a close to a 3:1 ratio; (2) different competing plans adopt rather disparate pass through rates, see Figure 1.1 ibid in Chapter 1. In any case, the above regression approaches are intrinsically confined to evaluating *level* changes in the existing exogenous capitation scheme.

Based on a full characterization of the equilibrium behavior of a broad class of price competition games in Federgruen and Lu [2016a], the objective of this paper is to provide the first estimates, under the various reform plans, of the above performance measures by *computing* the changes in the price equilibria and market shares. This in contrast to the unrealistic assumption that no bid changes would occur, the current system, as opposed to the above regression approach or, in the case of level changes within the current system, as opposed to the above regression approaches. We have applied our methodology to various consumer choice models whose parameters fit the year 2010's data for all 2478 counties with two or more private plans, offered in conjunction with the traditional FFS plan. Accounting for changes in price equilibria has a major impact on various performance measures. For example, if all subsidies were determined as the second lowest bid, The Congressional Budget Office estimates a cost saving for Medicare of 11% while our estimates range between 16.5% and 21%.

Medicare Advantage has allowed private insurance companies to offer private plans, as an alternative to the traditional Medicare option, which continues to be run by the Federal government. In 2003, private MA plans captured only 13% of the potential market; however, their share has steadily grown to 31% in 2015, see Figure 2.1.

In the MA program, the government pays most of the insurance premium of the different plans that are available to the beneficiaries. Currently, the federal government announces, each year, a county-specific capitation rate or premium subsidy[1]. In response to these pre-announced capitation

[1]Individuals have a specific risk score based on their prior medical history. This score is an estimate of the individual's expected covered costs, expressed as a multiple of the average individual's cost. The actual subsidy received for any given individual is determined as the "normalized' capitation rate , multiplied with the individual's risk score.

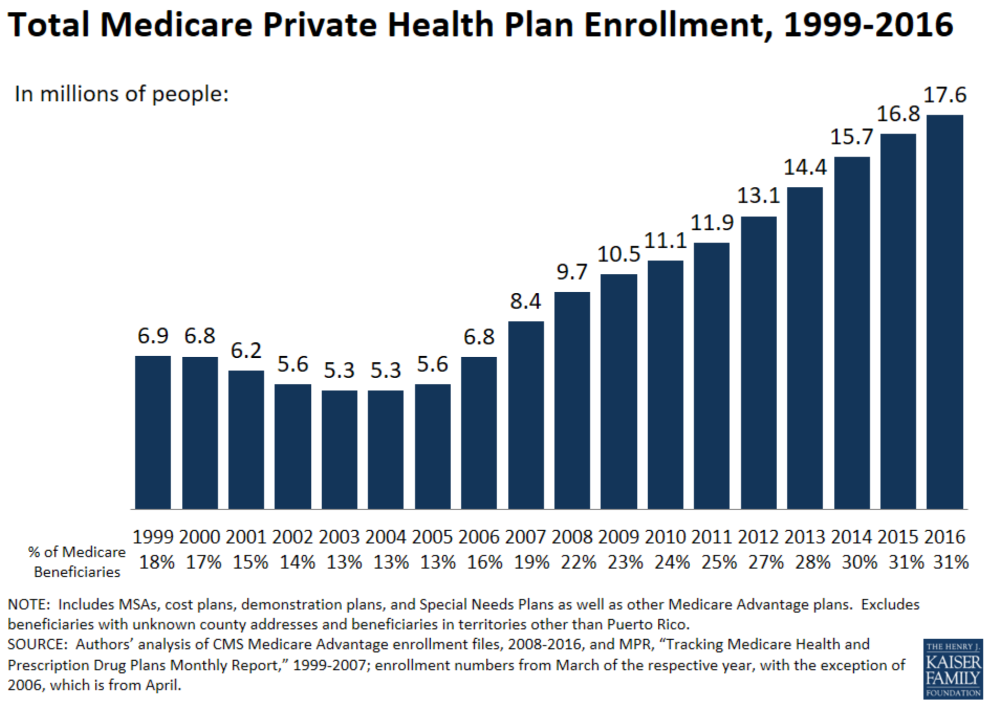



Figure 2.1: Total Private MA Enrollment

rates, insurance companies submit, each year, to the Center for Medicare & Medicaid Services (CMS), by a given deadline, one or several plans, each covering one or a collection of counties. For a plan to be eligible it must satisfy various criteria: In particular, it must provide benefits that are actuarially at least equivalent to those in the traditional Medicare plan, even though the specific provider network, menu of services and devices covered, as well as any associated copayments, etcetera, may be varied freely. All individuals covered by Medicare pay a monthly base premium, independent of their plan choice. When choosing a private plan, a beneficiary pays an additional premium, given by the *net* premium = nominal premium - capitation rate, when positive, or she receives a 75% rebate when the relative premium is negative. (In some cases, the rebate is offered in the form of additional benefits, as apposed to a cash rebate.) In an open enrollment period, beneficiaries choose one of the available alternative plans, i.e., the traditional Medicare plan or one of the private MA plans, with full knowledge of the associated net premia and rebates.

In an attempt to reduce the staggering costs of the Medicare program, there have been continuous proposals to alter the capitation levels. Some of these proposals have been legislated, e.g., the 2000 Benefits Improvement and Protection Act (BiPA) and, more recently, the 2011 Affordable

Care Act (ACA), often referred to as the Obama-care Act. Much more boldly, in 2011, Senator Wyden and then House Budget Committee Chairman Ryan advanced a proposal, the so-called Wyden-Ryan plan (W-R), to modify the competitive bidding process to one in which the capitation rate is no longer *exogenously* specified and pre-announced, but endogenously determined by the *second lowest* bid[2]. Those adopting the lowest premium plan would receive a rebate in the amount of 75% of the difference between the second lowest and the lowest bid. A similar proposal had been offered, in 2010, by the The Bipartisan Policy Center's Debt Reduction Task Force, chaired by Senator Pete Domenici and Alice Rivlin, the former Director of Office of Management and Budget in the Clinton Administration. The W-R plan became a focal point of the 2012 presidential campaign by the Romney-Ryan ticket, after Democratic senator Wyden had withdrawn his support.

In parallel, the W-R plan was widely discussed in the medical literature. For example, The *New England Journal of Medicine* published two articles, Antos [2012] and Aaron and Frakt [2012], presenting diametrically opposing arguments regarding the merits of the competitive bidding proposal, neither one with any formal analysis. More recently, the Congressional Budget Office (2013) evaluated, as an alternative to the second lowest bid scheme, one which specifies the subsidy as a weighted average of the premium bids (with the past year's market shares as the weights). Finally, in 2015, the Center for Health Policy at the Brookings Institute held a conference with the title "Strengthening Medicare for 2030". Rivlin and Daniel [2015] argued there that "changes to Medicare as we know it, are necessary to ensure that this popular, successful program is able to deliver high quality care at sustainable cost to the much larger population of older beneficiaries who will be eligible by 2030". Rivlin and Daniel focus on the same structural change in the Medicare insurance market, where the exogenous pre-announced capitation rate is replaced by either the second lowest bid or a weighted average thereof. The authors distinguish between "*Plan One*" where only the private MA plans would be part of the bidding process, and "*Plan Two*" where all plans, the traditional Medicare plan included, would do so.

As mentioned, the objective of this paper is to develop realistic estimates of the impacts any of the above reform plans would have on various performance measures of national interest, in particular: the cost savings to the Medicare program, out-of-pocket expenses for the beneficiaries,

[2] The actual formula is somewhat more complex in that it specifies the minimum of the second lowest bid and the traditional Medicare nominal premium. However, the latter is rarely lower than the second lowest bid.

the market shares of the traditional Medicare and the MA programs, as well as a social aggregate welfare measure. Thus far, all studies of these structural changes from an exogenous to an endogenously determined subsidy have been based on the restrictive assumption that all plans would continue to be offered at the same nominal premia and that their market shares would not be affected either. For example, the JAMA article by Song *et al.* [2012a] has estimated how much beneficiaries would have to pay in 2009, under the W-R plan, were they to stay with the traditional Medicare plan. (They also estimate that, on average, a beneficiary would have paid an additional $64, monthly, this represents 9% of the cost of the plan[3].) Based on the same assumption, the Congressional Budget Office (2013) estimated that the second-lowest bid option [weighted average bid option] would have resulted in an 11% [4%] reduction of federal spending, while increasing payments by affected beneficiaries by 11% [-6%].

In reality, of course, a change in the subsidy scheme, would result in significant changes in the nominal premium bids and market shares for the various plans, as firms would adjust their prices to the new rules and competitive dynamics, and beneficiaries would react to different net premia.

This fact is, of course, recognized by the public health policy makers. In their above mentioned position paper, Rivlin and Daniel [2015] wrote: "These [Song *et al.* 2012a] calculations suggest that competitive bidding in Medicare Advantage could produce substantial cost-savings under either method of defining the benchmark. However, our analysis and other studies assume that Medicare Advantage plans will submit bids to the CMS according to their historical tendencies. Unfortunately, it is hard to know what the MA bids would be under a new system not anchored to the cost of FFS Medicare."

The objective of this paper is to show how such estimates can, in fact, be derived from an appropriately estimated or calibrated oligopoly model for the county-by-county Medicare markets. Indeed, when incorporating the impact of competitive dynamics, we conclude that the government's cost savings associated with an endogenous capitation rate determined as the second lowest bid or

[3]Much larger estimates of annual costs of $6000 or more, for those reaching the age of 65 in the year 2030, have been propagated by the Obama campaign. However these numbers are based on a different provision in the W-R plan, where any given year's capitation rate, in any county, is capped at the prior year's value multiplied by the growth rate of the GDP plus one percentage point. These estimates assume that traditional Medicare costs will continue to grow at a rate significantly in excess of the growth rate of the country's GDP. In any case, the impact of this provision in the W-R plan is beyond the scope of this paper.

a weighted average of the bids, are even larger than what was estimated in Song *et al.* [2012a] and other such studies, and by a significant magnitude, indeed! For example, we estimate that, under the second lowest bid scheme, the cost savings would be between 16.5% and 21%, if "Plan Two" were adopted, and 8% under "Plan One".

Structural estimation of the changes in premium bids and market shares, requires one to (i) develop and estimate a competition model based on an underlying consumer choice model, (ii) characterize its equilibrium behavior and (iii) be able to compute the Nash-equilibrium or equilibria under the various subsidy schemes under consideration.

As to the first challenge, i.e., the estimation of a competition model for the Medicare market, such models were estimated by several authors, in particular Dowd *et al.* [2003], Hall [2007], Lustig [2008] and Nosal [2012]. All of these employ a MultiNomialLogit Model (MNL) or a variant thereof. However, all of the above models specify the plans' utility measure as *linear* in the nominal premium. As explained, in the Medicare markets, the consumer is only affected by the *net premium*, i.e., the difference between the nominal premium and the (very large) subsidy and her out-of-pocket costs or gains depend in a *non-linear* fashion on this net premium: recall that the beneficiary pays the full net premium, when positive, but receives, as a rebate, only 75% of (the absolute value of) the net premium, when negative. Curto *et al.* [2015] therefore specify and estimate a MNL model in which the consumer's utility measure depends on the net price in accordance with the above piecewise linear function. These authors employ detailed county-by-county data pertaining to the years 2006-2010. (The authors also estimate a *nested* MNL model, assuming beneficiaries first decide whether to enroll in the traditional Medicare plan, or one of the private MA plans. When choosing for the latter alternative, they make a second choice for a specific MA plan based on a second MNL model.)

After identifying an appropriate MNL-type competition model for the various county-by-county Medicare markets, its equilibrium behavior needs to be characterized and the equilibria computed, for each of the considered subsidy schemes. Here, we rely on results in Federgruen and Lu [2016a] covering a broad class of price competition models, with many diverse applications, in which the products' utility measure depends on its relative price in accordance with a general non-linear response function.

Part of our analysis is based on estimated models in Curto *et al.* [2015]; a second set of estimates is generated on the basis of MNL-competition models that are *calibrated* to the actual county-by-county data, for the year 2010. In a pure MNL-model, *all* beneficiaries in a given county are assumed to assign a utility value to each of the available plans that is drawn from a *single, common* distribution. This implies, for example, that those who in the past were enrolled in the traditional Medicare plan are just as likely to (re-)enroll there as the beneficiaries who had adopted a private MA plan. In reality, there is much inertia in these plan choices, as there is in many other markets, a phenomenon addressed and documented in Nosal [2012]. See Dube *et al.* [2009] for a treatment of inertia in general marketing models.

We therefore evaluate how our estimates vary when a switching cost is added to the utility measures. This requires segmenting the market based on the prior year's plan choice by the beneficiaries, and gives rise to a so-called Mixed MultiNomialLogit model (MMNL), again with a complex dependence of the utility measures on the plans' relative premia.

Here are some of our key findings: based on the 2010 calibrated competition models for the various counties, we observe a *significant reduction* of the equilibrium premia, compared to those selected under the prevailing, exogenously specified capitation rates. As a consequence, we estimate that the W-R plan would result in a reduction of approximately 18.5% in the capitation rates and of 16.5% in the government's costs. Based on the results for the estimated model, the decrease in the government costs would be 21%. A 16.5% reduction in the Medicare budget would have saved close to $80 billion in the 2012 calendar year alone; this, compared with a total of $68 billion from 2012–2016, predicted to be saved under the Affordable Care Act, due to its mandated reduction of capitation rates. For beneficiaries continuing to opt for the traditional Medicare plan, the average monthly cost is roughly $64, comparable to those estimated in Song *et al.* [2012a], under the assumption of unaltered premia and market shares.

The remainder of the paper is organized as follows. Section 2.2 provides a review of the related literature. Section 2.3 describes the relevant institutional facts about the Medicare Market. Section 2.4 describes the competition model and its equilibrium behavior. Section 2.5 and 2.6 develop estimates based on calibrated vs estimated competition models, respectively. Section 2.7 provides a concluding summary.

## 2.2 Literature Review

Our paper is a contribution to the recent literature on competition models for the Medicare industry. As mentioned, the total subsidy cost in this industry, as covered by the Medicare budget, exceeded $600 billion in 2014; this is approximately three times the total revenues in the US hotel industry or the pharmacy and drug sales industry.

As mentioned in the Introduction, most of the existing competition models for the Medicare industry have used a MNL model as the underlying consumer choice model. With the exception of the recent paper by Curto *et al.* [2015], these models have specified the utility measure for the competing insurance plans as *linear* functions of the *nominal* premium. We have explained that both of these structural assumptions fail to be satisfied in the Medicare market where consumer choices are only affected by the *net* premium, i.e., the difference between the nominal premium and the government subsidy, or capitation rate. Moreover, out-of-pocket costs are a non-linear function of the net premium. Examples include Dowd *et al.* [2003], Hall [2007], Lustig [2008] and Nosal [2012]. Some of these models treat some of the coefficients as random or introduce other segmentations of the population of eligible beneficiaries. In contrast, Curto *et al.* [2015] specify each plan's utility measure as a piecewise linear function of its *net* premium.

Curto *et al.* [2015] use their model to assess the impacts of level changes in the current capitation system. To this end, they estimate a linear regression equation to predict how premium bids respond to changes in the (exogenous) capitation level. They also estimate a consumer choice model for these markets, enabling them to predict how the plans market shares, in turn, would respond to the premium changes. In contrast our approach is geared to estimating the impacts of general premium support schemes, not just level changes within the current structure; moreover, we obtain all of our estimates, consistently, from computed equilibria within the same competition model.

In this paper, we employ the same structural assumptions as Curto *et al.* [2015], directly anchored on the Medicare payment structure. Depending upon the choice of the subsidy scheme, this implies that each plan's utility measure is, in general, a complex function of *all* nominal premia in the market. Even if the subsidy or capitation rate is an exogenously given constant, as in the current Medicare structure, the fact that out-of-pocket costs or rebates depend on the net price via a non-linear function, creates significant complications when attempting to characterize the equilibrium behavior of the competition model.

Several papers have addressed he equilibrium behavior in MNL-type models, but all under the assumption that the utility measures' price dependence is confined to a dependence on the the product's own and absolute price. It is well known that an equilibrium exists, in the standard MNL model, see Anderson *et al.* [2001], Bernstein and Federgruen [2004] and Gallego *et al.* [2006]. However, an equilibrium may fail to exist in various generalizations of the basic MNL model, for example Mixed MNL models (MMNL) where the market is segmented and the structure of the utility functions varies by segment.(In the latter case, Allon *et al.* [2013] have shown that an equilibrium may fail to exist while providing specific market share conditions under which the existence question can be answered in the affirmative.) Liu [2006], Li and Huh [2011], and Gallego and Wang [2014] study various conditions under which an equilibrium is guaranteed to exist in a nested logit model.

Federgruen and Lu [2016a] have characterized the equilibrium behavior in general MNL models, in which the products' utility measure depends on its *relative* rather than its *absolute* price and this according to a general non-linear response function and various reference value schemes. The authors document that this structure is pervasive in many marketing, operations management and prospect theoretical industrial organization models. The analysis in this paper makes explicit use of several of the results in Federgruen and Lu [2016a].

Another recent stream of papers address how insurance firms adjust their nominal premia in response to changes to the pre-specified capitation rate levels. More specifically, several prominent recent studies, e.g., Song *et al.* [2012b], Song *et al.* [2013], Cabral *et al.* [2014], Duggan *et al.* [2014] and Curto *et al.* [2015] have applied regression models to estimate the *pass-through rates*, i.e., the absolute change in the premia due to changes in the capitation rates. These studies reach rather different conclusions, with estimated pass through rates varying between 37% and 100% across the different studies, leaving policy makers with a great deal of ambiguity. Moreover, these reduced form approaches assume a specific structural form, for example linear or log-linear, for the dependency of the firms' premia on the subsidy (capitation) levels, which may not be consistent with any plausible underlying price competition models. As an alternative to these *reduced* form approaches, we employ a structural estimation approach; more specifically, we conduct counterfactual studies within the above estimated or calibrated competition models for the Medicare markets, to estimate these pass-through rates.

In Section 1.1, we cited Song *et al.* [2012a] and Congressional Budget Office (2013) as two studies attempting to estimate the impacts of reform plans which replace the constant *exogenous* capitation rate with one that is *endogenously* determined (e.g., the second lowest bid or a weighted average of the bids.) As mentioned, these studies assume that the firms' nominal premia and market shares are not affected by this structural change.

Our paper may also be viewed as a contribution to the emerging operations research and economics literature employing game-theoretical models to understand the impact of competitive schemes on the performance of (specific segments of) the healthcare industry. See Lu and Donaldson [2000] for a survey of the 20th century economics literature on performance-based contracting. Fuloria and Zenios [2001] develop a dynamic principal-agent model to determine contract terms for providers and purchasers contingent on observed outcomes, such as mortality and medical complications. So and Tang [2000] constructed a Medicare contract for the reimbursement of drug prescriptions with a clinical outcome-based performance metric. Lee and Zenios [2012] studied evidence-based incentive systems in the context of dialysis treatment for patients with end-stage renal disease. Jiang *et al.* [2012] uses a principal-agent game-theoretical model to build performance-based contracts for outpatient medical service.

## 2.3 The Medicare Market

Most beneficiaries in the Medicare program face the choice between the government run traditional Medicare program or any one of a set of private Medicare Advantage plans. The traditional program allows its enrollees to select their providers, at their will, among all those accepting this Medicare insurance program. No premium is charged beyond the base premium charged to all beneficiaries, regardless of their plan choice. For example, in 2014, the base premium per individual was $104.90 per month, assuming the individual was married with an annual income below $170,000.

The private MA plans are typically offered by Health Maintenance Organizations (HMO) or Preferred Provider Organizations (PPO). Except for certain emergencies, a HMO usually limits coverage to health care providers who work for or contract with the HMO. A PPO also specifies a network of providers but affords its members the option to select providers from outside the network. However, when an enrollee of a PPO plan uses an out-of-network provider, she is reimbursed at a lower rate and may face higher deductables and co-payments. The MA plans differentiate themselves along various dimensions: first, they charge a (positive or negative) net premium to each enrollee, above and beyond the base premium that is charged to *all* Medicare beneficiaries. This net premium (or rebate) was mentioned in the Introduction, and is explained in more detail below. In addition, MA plans may offer additional benefits (vision, dental coverage) beyond those offered by the traditional Medicare program. The plans offer different provider networks, have different quality of service standards, co-payments, etcetera.

In addition, there are often several plans known as "private-fee-for-service" (PFFS) plans. These plans offer terms that are identical or virtually identical to the traditional government-run Medicare program. While somewhat prevalent at the start of the century, their market share has dropped to 7% in 2011. For all these reasons, we aggregate these PFFS plans with the traditional government-run FFS plans.

The private MA insurance companies submit plans to the CMS with a nominal premium bid. When approved by the CMS, the plan is offered to the beneficiaries in the relevant county, along with all other approved MA plans and the traditional Medicare program. The Medicare program subsidizes a large part of the nominal premium by establishing a uniform, though county-specific, capitation rate. A beneficiary is thus affected only by the *net* premium, i.e., difference between the nominal premium and the capitation rate. When the net premium is *positive*, it is paid, in full, by the beneficiary, on top of the above mentioned base premium charged to all Medicare participants. When the net premium is *negative*, 75% of its absolute value is given to the beneficiary as a rebate; alternatively, the insurance plan may provide other benefits whose momentary value is equivalent to the rebate.

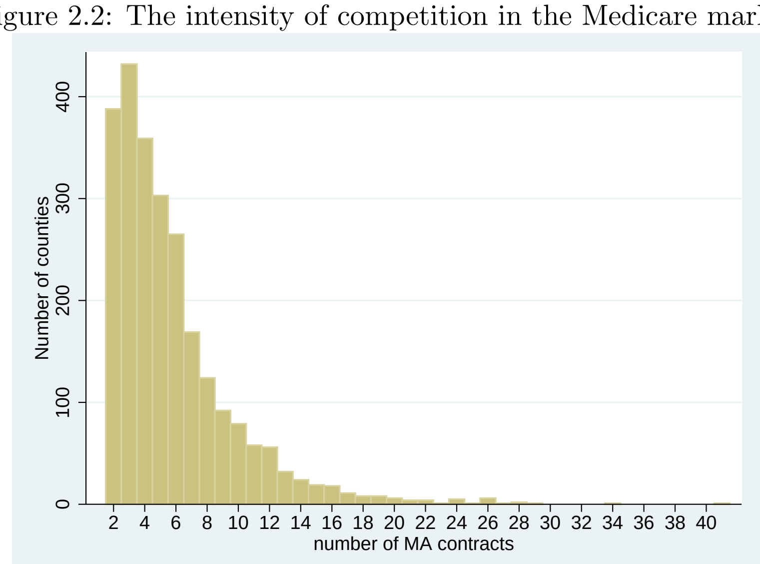
As mentioned the capitation rate is currently *exogenously* specified by CMS and pre-announced to the insurance companies, before they submit their bids. The capitation rate is the amount paid by the government to the insurance company for a (normalized) individual with a risk score of one. The risk score is meant to predict by what factor an individual's expected annual medical costs exceeds that of the average beneficiary. As a consequence, an MA insurance plan covering an arbitrary individual is paid the (normalized) capitation rate multiplied by the individual's risk score.

Since 2007, CMS employs a risk score determined by the so-called "Hierarchical Condition

Categories" (HCC) model. The score is based on demographic data, as well as the individual's disease record, employing seventy disease - category indicator variables, themselves distilled from roughly 15000 ICD-9 codes that providers list on claims. The predictive accuracy of the HCC risk score is actively debated and analyzed, see e.g. Brown *et al.* [2014] and the references therein. However, as in all prior models of the Medicare market, we assume that the risk score is an unbiased estimator of the actual cost incurred by the insurance plans, relative to that of a normalized beneficiary with a risk score of one. Figure 2 in Curto *et al.* [2015] shows that the market share of MA plans decreases modestly with the risk score, suggesting that the private MA plans are disproportionately desired by healthier individuals.

There is abundant evidence that the Medicare market in a typical county, is heavily concentrated. This is an important feature in the context of our paper's objectives since it demonstrates that changes in the subsidy structure or level can be expected to result in significant changes of the premium choices of the competing MA plans as well as their market shares, contrary to what has been assumed in existing studies, e.g., Song *et al.* [2012b] and Feldman *et al.* [2012].

Figure 2.2: The intensity of competition in the Medicare market




First, Curto *et al.* [2015] report that, between 2006 and 2011, the average beneficiary had a choice of some 4 MA plans, typically 3 HMOs and one PPO. (It should be noted that the number of competing MA plans varies greatly by county. Figure 2.2 exhibits, for the 2010 calendar year, the number of counties with a given number of MA contracts. ) The same authors also point out that in more than 75% of US counties, the *three* largest plans have more than 90% of the MA market. In close to half of the counties the two largest plans. attract more than 90% of the total MA market. Industry concentration tends to be higher in rural as opposed to the urban areas. See Table A.2 in Curto *et al.* [2015]. Indeed, in the years 2006-2011, the Herfindahl-Hirschman Index (HHI) averaged 0.477 for urban areas and 0.547 for rural areas. (The HHI is defined as the sum of the squares of the individual firms' market shares.) The US Department of Justice, in its 2010 Horizontal Merger Guidelines characterized any market with an HHI value of 0.25 or higher as "highly concentrated".

Finally, a statistical multi-year study by Song *et al.* [2012b] shows that the insurance companies adjust their premia to changes in the capitation rates, even though the latter are only partially correlated with the actual cost rates per beneficiary. Indeed, their regression model, allowing for many potential explanatory variables, identifies only the capitation rate, and the number of competitors in the market as having a statistically significant impact on the premia. The fact, that the premia decrease significantly as a function of the number of competing plans, is further evidence for the fact that the market is a concentrated oligopoly with imperfect competition.

As in prior studies, we treat every county as a separate market. This is a slight simplification, since insurance companies often offer the same plan in a "service area" consisting of several counties. If so, they select a single (nominal) bid that gets adjusted at the county level based on how the county capitation rate compares with the average capitation rate in the service area. Like Curto *et al.* [2015], for example, we continue to view each county as a separate market because insurers have the option of specifying "more granular" plans and differentiating their bids on a county-by-county basis.

Finally, the Medicare industry differentiates between "contracts" and "plans". An MA insurance company submits a contract which may consist of several plans, each with slightly differentiated terms and (nominal) premium bids. In our analysis, we represent each county-contract pair as a single plan whose attributes are determined as a weighted average of the attributes of the underlying plans, and with a market share given by the aggregate of the shares of the individual plans. This is hardly an approximation, as typically a single plan in a multi-plan contract is adopted by the vast majority of those enrolled in the contract. With this aggregation step, each county is represented as an oligopoly competition model in which each competitor offers a single "representative plan".

This representation is identical to that adopted in almost all of the recent literature, geared towards an estimation of the demand functions in the Medicare market, see in particular Hall [2007] and Nosal [2012]. (Curto *et al.* 2015 is an exception estimating on a plan-level basis.)

## 2.4 The Competition Model and Its Equilibrium Behavior

Consider a given county in a given calendar year. There are $N$ competing private insurance plans which are part of the Medicare Advantage Program, along with the traditional government owned Medicare program. We use the index $i$ to differentiate among the plans and assign index $i = 0$ to the traditional Medicare program, as mentioned with a declaiming, but still dominant market share of 61%. The plans are differentiated by their *nominal* or absolute monthly premium, as well as various coverage attributes, including the specific provider network, menu of services and devices covered, as well as any associated copayments, etcetera.

Both in the existing Medicare program, and in any of its reform proposals, a major portion of the beneficiary's premium is covered by a government subsidy or *capitation rate*. Thus, potential customer react to *relative* premia, defined as the difference between the nominal premium and the subsidy. The subsidy may be an *exogenously* specified, and pre-announced constant value, as in the existing Medicare program. Alternatively, it may be specified *endogenously* as a function of the set of nominal premia in the market; this is the core idea in many of the Medicare reform proposals, see the Introduction. Of particular interest in these proposals, is a specification of the subsidy as the lowest, second lowest or a weighted average of the nominal bids.

Each individual beneficiary is assigned a risk score based on her prior medical history. The subsidy provided by the federal government is the "base" capitation value multiplied by the individual's risk score. The cost structure of any given plan is clearly affine in the total number of beneficiaries covered; aside from administrative costs, that, in the short run, can be treated as fixed, the plans' layouts are the sum of the layouts for the various beneficiaries covered by the plan and therefore, in expectation, strictly proportional to the number of individuals covered.

We employ the following notations where the index $i$ covers the range $i = 0, 1, 2, \ldots, N$:

$P$ = the size of the population of Medicare beneficiaries in the county;

$c_i$ = the expected cost incurred by insurance plan $i$, for a "normalized" beneficiary, with a risk score of one;

$b_i$ = the bid submitted by plan $i$;

$g(\mathbf{b})$ = the subsidy or capitation rate paid by the Medicare program for any MA-enrolled beneficiary, with a risk score of one;

$p_i$ = the *net* monthly premium charged by plan $i$, i.e., $p_i = b_i - g(\mathbf{b})$

$d_i$ = the expected number of beneficiaries in the county who choose plan $i$;

$\pi_i(\mathbf{b})$ = the expected profit for plan $i$.

To cover both the existing Medicare structure as well as those embedded in the various reform proposals, under discussion, we provide a unified treatment by specifying the subsidy or capitation rate as a general function $g(\cdot)$ of the vector of bid premia $\mathbf{b}$. As we will note, the qualitative properties of the equilibrium behavior and, in particular, the equilibrium prices themselves depend significantly on the specification of the function $g(\cdot)$.

The following set of specifications covers the spectrum of proposals:

$$g(\mathbf{b}) = \begin{cases} C, & \text{exogenous and constant} \quad (2.1a) \\ b_{(1)}, & \text{the lowest bid} \quad (2.1b) \\ b_{(n)}, & \text{the } n\text{-th lowest bid, with } n \geq 2 \quad (2.1c) \\ \sum_j w_j b_j, & \text{a weighted average of selected bids} \quad (2.1d) \end{cases}$$

Structure (2.1a) underlies the common Medicare system, while (2.1c) and (2.1d) have been proposed in the original Domenici-Rivlin and Wyden-Ryan plan, the 2013 Congress Budget Office's variant thereof as well as the proposals in Rivlin and Daniel [2015]. (Thus far, no one appears to have proposed that the subsidy be specified as the *lowest* bid; we model this option, however, as there is no good reason for its exclusion, other than the perhaps politically inconvenient fact that no beneficiaries receive a rebate, under this structure.)

Each beneficiary $j$ assigns a utility value to each of the $N$ available private MA plans, specified as follows

$$u_{ij} = a_i - \gamma_i \cdot f\left(b_i - g(\mathbf{b})\right) + \epsilon_{ij}, \quad i = 1, 2, \ldots, N, \; j = 1, 2, \ldots. \tag{2.2}$$

The intercept $a_i$ denotes the aggregate impact of all of plan $i$'s observable non-price attributes. Different papers have adopted different specifications for these intercepts. For example, in Curto *et al.* [2015], $a_i = x_i^T \eta$, with $x_i$ an M-dimensional vector of plan characteristics. These include indicator variables for each possible quality score (e.g., 3.5 stars or 5 stars), indicator variables expressing whether the plan is associated with supplemental benefits (for example, vision or dental coverage) or whether it is bundled with Part D (i.e., prescription drug) benefits. See Dowd *et al.* [2003], Hall [2007], Lustig [2008] and Nosal [2012] for alternative specifications of the intercept values of $\{a_i, i = 1, 2, \ldots, N\}$.

In the existing Medicare structure, the response function $f(\cdot)$ is a (convex) piecewise linear function of the form:

$$f(x) = \begin{cases} \alpha x, & \text{if } x > 0 \\ \beta x, & \text{if } x \leq 0 \end{cases}, \quad \text{with } 0 < \beta < \alpha. \tag{2.3}$$

(Indeed, in the existing Medicare structure and all reform proposals we are aware of, $\alpha = 1$ and $\beta = 0.75$. However, the rebate fraction $\beta$ may itself be a parameter in the ultimate restructuring of the Medicare program; we therefore maintain general parameters $\alpha$ and $\beta$, as in (2.3).) All of the results in this paper are anchored on the piecewise linear structure of the response function $f(\cdot)$, however, extensions to general increasing functions $f(\cdot)$ are entirely feasible given the results in Federgruen and Lu [2016a].

The last term $\epsilon_{ij}$ in (2.2) represents a random unobserved component of the customer $j$'s utility for plan $i$, which varies by customer. These random variables $\{\epsilon_{ij}\}$ are assumed to be i.i.d across plans and beneficiaries, following a standard type 1–extreme value or Gumbel distribution, i.e., $\mathbb{P}\left(\epsilon_{ij} \leq x\right) = \exp\left(-\exp(-x + \delta)\right)$ where $\delta$ is Euler's constant (0.5772).

The specification of the utility measure in (2.2) implies that all beneficiaries "draw" a utility value for any given plan, from the same *common* distribution. However, the population of beneficiaries may need to be partitioned into several segments, each with its own specification of the parameter vector and price sensitivity coefficient $\{\gamma_i\}$. Inertia and switching costs, may need to be added to reflect that the odds of a beneficiary choosing plan $j = 1, 2, \ldots, N$ is significantly larger among prior enrollees of this plan, as compared to those that were enrolled in a different plan. Similarly, the weight attributed to various plan characteristics may depend on the beneficiary's risk score, necessitating a segmentation according to this risk score measure. Finally, as mentioned, some of the above mentioned papers treat part of the parameters in (2.3) as random, resulting in another type of segmentation and a so-called Mixed MultiNomialLogit consumer choice model.

Our base model and analysis are anchored on the (unsegmented) specification in (2.2), however, in Section 2.4.1, we pursue a segmented representation, that incorporates the above switching costs.

Since all US residents, of age 65 and older, are eligible for Medicare coverage, the traditional Medicare program functions as the "outside good", We specify its utility measure:

$$u_{0j} = a_0 - \gamma_0 f\left(b_0 - g(\mathbf{b})\right) + \epsilon_{0j}, \quad j = 1, 2, \ldots, \tag{2.4}$$

In the current Medicare structure, enrollees in the traditional program pay nothing (beyond the base premium charged to all beneficiaries.). Several of the reform proposals maintain this feature: Rivlin and Daniel [2015] refer to these a Plan I-proposals. Others, for example, the original W-R plan, have the traditional plan participate in the competitive bidding process. Both types can be accommodated by appropriate choices of the price sensitivity coefficients $\{\gamma_i, i = 0, 1, 2, \ldots, N\}$.

$$\underline{\text{Plan I:}} \quad \gamma_0 = 0, \gamma_1 = \gamma_2 = \cdots = \gamma_N = \gamma.$$

$$\underline{\text{Plan II:}} \quad \gamma_0 = \gamma_1 = \gamma_2 = \cdots = \gamma_N = \gamma.$$

In (2.4), the $\{\epsilon_{0j}\}$-term again represents random unobserved terms assumed to be i.i.d with the same distribution as the noise terms in (2.2).

Since our consumer choice model is of the MNL-type, we obtain the following well-known expression for the *expected* number of beneficiaries that are enrolled in each plan $i = 0, 1, \ldots, N$:

$$d_i(\mathbf{b}) = P \cdot \frac{\exp\left(a_i - \gamma_i \cdot f\left(b_i - g(\mathbf{b})\right)\right)}{\sum_{j=0}^{N} \exp\left(a_j - \gamma_j \cdot f\left(b_j - g(\mathbf{b})\right)\right)}, \tag{2.5}$$

see for e.g., Anderson *et al.* [2001]. We thus obtain the following expression for the expected profit earned by the various plans.

**Lemma 4** *Plan i's profit is given by*

$$\pi_i(\mathbf{b}) = (b_i - c_i) d_i(\mathbf{b}), \quad i = 0, 1, \ldots N. \tag{2.6}$$

**Proof.** Let $D_i(\mathbf{b})$ denote the number of county beneficiaries that enroll in plan $i$, with $d_i(\mathbf{b})$ as its expected value. Plan $i$'s profit value is given by

$$\pi_i(\mathbf{b}) = \sum_{l=1}^{D_i(\mathbf{b})} \left(r_l g((b)) + p_i - r_l c_i\right),$$

with $r_l$ the random risk score for individual $l$, since Medicare's subsidy for individual $l$ to insurance plan $i$ is given by $r_l g(\mathbf{b})$ and the expected costs incurred for this beneficiary is $r_l c_i$. Since $\mathbb{E}[r_l] = 1$ and $D_i(\mathbf{b})$ is independent of $\{r_l\}$, the conditions for Wald's Lemma are satisfied and, thus, one obtains (2.6). □

The expected profit functions thus satisfy the general structure analyzed in Federgruen and Lu [2016a]. We review key equilibrium results derived there, as needed in our subsequent analysis:

**Proposition 5** (Existence of a pure Nash equilibrium) *A pure Nash Equilibrium exists under any of the following structures:*

*(i) an exogenous constant capitation rate, see* (2.1a).

*(ii) a capitation rate given by the lowest or n-th lowest bid and* $\gamma_0 = \gamma_1 = \cdots = \gamma_N$ *(Plan II).*

*(iii) a capitation rate given by the lowest bid and* $\gamma_0 = 0 < \gamma_1 = \gamma_2 = \cdots = \gamma_N$ *(Plan I), under the following conditions*

$$\sum_{j \neq i} \frac{\partial_- d_i}{\partial b_j}(b^{-i}_{(1)}, b_{-i}) \leq \left|\frac{\partial_+ d_i}{\partial b_i}(b^{-i}_{(1)}, b_{-i})\right|, \qquad (\tilde{D})$$

$$(b_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial b_j} \text{ is quasi-convex in } b_i \in [c_i, b^{-i}_{(1)}). \qquad (M)$$

*(iv) a capitation rate given by the* $n-th$ *lowest bid* $(n \geq 2)$ *and* $\gamma_0 = 0 < \gamma_1 = \gamma_2 = \cdots = \gamma_N$ *(Plan I), under the following conditions:*

$$\left|\frac{\partial_+ d_i}{\partial p_i}(b^{-i}_{(n)}, b_{-i})\right| \geq \sum_{j \neq i} \frac{\partial_- d_i}{\partial b_j}(b^{-i}_{(n)}, b_{-i}), \qquad (D')$$

$$\sum_{j \neq i} \frac{\partial_+ d_i}{\partial b_j}(b^{-i}_{(n-1)}, b_{-i}) \geq \left|\frac{\partial_- d_i}{\partial b_i}(b^{-i}_{(n-1)}, b_{-i})\right|,$$

$$(b_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial b_j} \text{ is non-decreasing in } b_i \in (b^{-i}_{(n-1)}, b^{-i}_{(n)}), \qquad (M')$$

**Proof.** Part (i) follows from Theorem 1 in Federgruen and Lu [2016a], part (ii) from Theorem 4 and Lemma 2, there. Part (iii) and (iv) follows from Theorem 4 and Theorem 6, respectively. □

Condition $(D')$ is a variant of the classic dominant diagonal condition. This condition merely precludes that a uniform premium increase by all plans would result in an increase of any of the plans' enrollment volumes. The diagonal dominant condition goes back to Arrow *et al.* [1959] and Hadar [1965] and is a standard assumption in many price competition models, see e.g., Vives [2001], Bernstein and Federgruen [2004], Farahat and Perakis [2011] and Allon *et al.* [2013]. As innocuous as it appears, it may nevertheless be violated in special instances. Note that condition $(\tilde{D})$ is an (even) weaker variant of the dominant diagonal condition; it requires that any plan $i$'s enrollment volume does not increase when all plans increase their (nominal) premium by the same amount, but only when plan $i$'s premium is the lowest in the market.

The case where the capitation rate is determined as a weighted average of the premia, i.e., structure (2.1d), fails to be covered by Proposition 5. Indeed, under this subsidy structure, plans may cease to be substitutes, i.e., it is possible that a unilateral premium increase by a given plan $i$ causes the sales volume of some other plan $j \neq i$ to go down. Since the plans are no longer guaranteed to be substitutes, the existence of a (pure) Nash equilibrium can not be guaranteed either, a well known phenomenon, see e.g., Section 2.3.2 in Vives [2001] and Section 8 in Federgruen and Lu [2016a] for a more detailed discussion.

When the capitation rate is exogenously specified as a constant value, see (2.1a), as well as when it is endogenously determined as the *lowest* bid, it is, in fact, possible to prove that the competition model has the additional structure of (log-)super-modularity, see Vives [2001] and Topkis [1998]. (Log-)super-modularity implies that the *set* of pure Nash equilibrium, if not a singleton, is a lattice with a component-wise smallest and a component-wise largest element:

**Proposition 6** *The competition model is (log-)super-modular when*

*(i) the capitation rate is exogenously determined, see* (2.1a).

*(ii) the capitation rate is endogenously determined as the lowest bid in a Plan II-setting i.e., when* $\gamma_0 = \gamma_1 = \cdots = \gamma_N$.

**Proof.** (i) and (ii) follow from Theorem 1 and 5 in Federgruen and Lu [2016a], respectively. □

An additional implication of the price game being (log-)supermodular is the fact that an equilibrium may be computed with a simple tatônnement scheme: starting with an arbitrary price vector $b^{(0)}$, one iteratively computes a best response price for each of the $N$ firms to the most recently generated prices of the competitors. The scheme is guaranteed to converge to an equilibrium. Moreover, when the tatônnement scheme is started at $b^{min}[b^{max}]$, it is guaranteed to converge to $\underline{b}^*[\bar{b}^*]$. It is therefore possible to unequivocally determine whether the game has a unique equilibrium by starting the tatônnement scheme, both at $b^{min}$ and at $b^{max}$, and checking whether the two schemes converge to the same limit point.

### 2.4.1 Competition Model with Switching Costs

The utility measures in (2.2) and (2.4) while random, are identically distributed. As explained above, there are several considerations that challenge this homogeneity assumption. As a consequence, the market may need to be partitioned into a set of $K \geq 2$ segments, each with its own population size $P_k$ and its own specification of the random utility measures in (2.2) and (2.4), giving rise to an MMNL model, as follows:

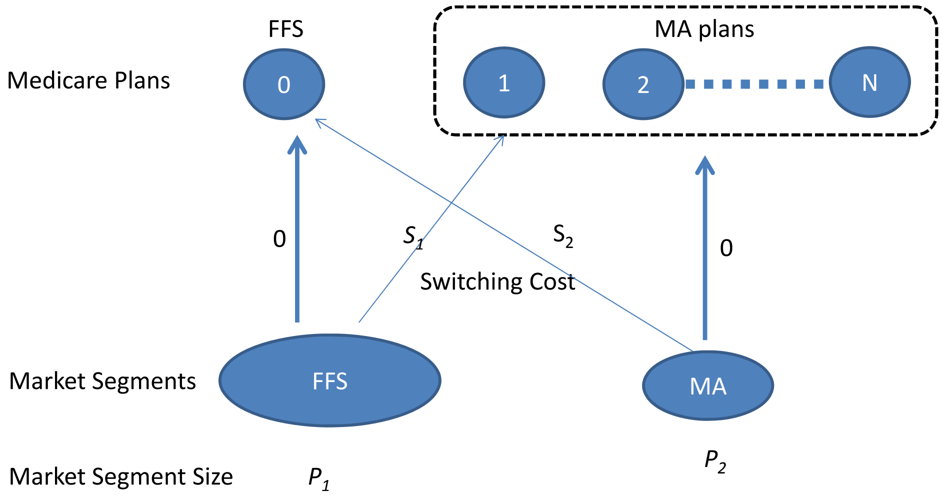
$$u_{ijk} = a_{ik} - \gamma_i f(b_i - g(b)) + \epsilon_{ijk}; \quad i = 1, 2, \ldots, N; \ k = 1, 2, \ldots, K; \ j = 1, \ldots \tag{2.7a}$$

$$u_{0jk} = a_{0k} - \gamma_0 f(b_0 - g(b)) + \epsilon_{0jk}; \quad k = 1, 2, \ldots, K; \ j = 1, \ldots \tag{2.7b}$$

Here $u_{ijk}$ denotes the utility attributed by the $j$-th potential customer in segment $k$ to plan $i$ $(k = 1, 2, \ldots, K; \ i = 0, 1, \ldots, N)$. $\{\epsilon_{ijk}\}$, again, represent a set of i.i.d. random variables with the Gumbel distribution. The MMNL model, without subsidization, i.e., with $g(b) \equiv 0$, has been employed in countless studies in industrial organization, marketing and operations management, among many other areas, see e.g. Caplin and Nalebuff [1991] and Allon *et al.* [2013]. Even in the base model, without subsidization $[g(b) = 0]$, a pure Nash equilibrium may fail to exist, see Allon *et al.* [2013]. The latter have developed a set of sufficient conditions for the existence of a pure Nash equilibrium or its uniqueness. It remains an open question how these conditions can be generalized for our model with a general subsidy structure $g(\cdot)$, specified as one of the cases (2.1a) - (2.1c).

As mentioned above, we have applied this model to incorporate the impact of switching costs. To this end, we partition the population of all beneficiaries in a given county into two segments: segment $k = 1$ represents beneficiaries, who are enrolled in the traditional Medicare program, in the prior year; segment $k = 2$ represents the remainder of the market. A switching cost $S_1$ is incurred when a beneficiary switches from the traditional Medicare program to a MA plan, and $S_2$ when the switch is in the opposite direction, see Figure 2.3 for a pictorial representation of the market segmentation and switching costs.

Figure 2.3: Price competition with switching cost




The general utility measures, thus, take the form,

$$u_{ij1} = a_i - S_1 \mathbb{I}(i \neq 0) - \gamma_i \cdot f\left(b_i - g(\mathbf{b})\right) + \epsilon_{ij1}; \quad i = 0, 1, 2, \ldots, N; \; j = 1, \ldots \tag{2.8}$$

$$u_{ij2} = a_i - S_2 \mathbb{I}(i = 0) - \gamma_i \cdot f\left(b_i - g(\mathbf{b})\right) + \epsilon_{ij2},; \quad i = 0, 1, 2, \ldots, N; \; j = 1, \ldots \tag{2.9}$$

For each MA plan $i = 1, \ldots, N$, its market share in segment $k = 1, 2$ is given by

$$\begin{aligned} d_{i1}(\mathbf{b}) &= P_1 \frac{\exp\left(a_i - \gamma_i \cdot f\left(b_i - g(\mathbf{b})\right) - S_1\right)}{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right)\right) + \sum_{k=1}^N \exp\left(a_k - \gamma_k \cdot f\left(b_k - g(\mathbf{b})\right) - S_1\right)}, \\ &= P_1 \frac{\exp\left(a_i - \gamma_i \cdot f\left(b_i - g(\mathbf{b})\right)\right)}{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right) + S_1\right) + \sum_{k=1}^N \exp\left(a_k - \gamma_k \cdot f\left(b_k - g(\mathbf{b})\right)\right)}, \quad (2.10) \\ d_{i2}(\mathbf{b}) &= P_2 \frac{\exp\left(a_i - \gamma_i \cdot f\left(b_i - g(\mathbf{b})\right)\right)}{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right) - S_2\right) + \sum_{k=1}^N \exp\left(a_k - \gamma_k \cdot f\left(b_k - g(\mathbf{b})\right)\right)}, \quad (2.11) \end{aligned}$$

Similarly, the market share of the FFS plan 0 in each segment is given by

$$\begin{aligned} d_{01}(\mathbf{b}) &= P_1 \frac{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right)\right)}{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right)\right) + \sum_{k=1}^N \exp\left(a_k - \gamma_k \cdot f\left(b_k - g(\mathbf{b})\right) - S_1\right)}, \\ &= P_1 \frac{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right) + S_1\right)}{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right) + S_1\right) + \sum_{k=1}^N \exp\left(a_k - \gamma_k \cdot f\left(b_k - g(\mathbf{b})\right)\right)}, \quad (2.12) \\ d_{02}(\mathbf{b}) &= P_2 \frac{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right) - S_2\right)}{\exp\left(a_0 - \gamma_0 \cdot f\left(b_0 - g(\mathbf{b})\right) - S_2\right) + \sum_{k=1}^N \exp\left(a_k - \gamma_k \cdot f\left(b_k - g(\mathbf{b})\right)\right)}, \quad (2.13) \end{aligned}$$

The total number of beneficiaries enrolled in plan $i$ is given by $d_i(\mathbf{b}) = d_{i1}(\mathbf{b}) + d_{i2}(\mathbf{b})$, and its profit function by $\pi_i(b_i, b_{-i}) = (b_i - c_i)d_i(\mathbf{b})$.

Without loss of generality, we normalize $S_2 = 0$, otherwise, if $S_2 \neq 0$, (2.10)–(2.13) may be normalized via the following transformations

$$\bar{a}_1 = a_1 - S_2 \quad \text{and} \quad \bar{S}_1 = S_1 + S_2.$$

We analyze the model for a given switching cost value $S_1$ and a given estimate $\hat{\gamma}$ for the price sensitivity coefficient $\gamma_i = \gamma$.

## 2.5 Estimating the Impact of Medicare Reform Proposals: A Calibration Model

We are now ready to pursue the main questions of our study, i.e., the estimation of the impact various reform plans would have on the principal performance measures of interest, in particular, (i) the cost savings to the Medicare program; (ii) the out-of-pocket expenses for the beneficiaries; (iii) the premium bids; (iv) the market shares of the traditional Medicare program and the MA program; as well as (v) an aggregate welfare measure.

In this section, we derive these estimates based on models of the type developed in Section 2.4.1, that are calibrated to publicly available county-by-county data for the calendar year 2010. As explained, we treat each county as a separate market. We start with the base model as specified by the demand volume and profit equations in (2.5) and (2.6) with $g(b) = C$ and $C$ the county's prevailing benchmark or capitation rate, in 2010 (for a normalized beneficiary with a risk score of one.) Furthermore, in the 2010 and current structure, the traditional Medicare program does not participate in the competitive bidding process; in other words, the existing structure is a Plan-I type and has $\gamma_1 = \gamma_2 = \cdots = \gamma_N = \gamma$ and $\gamma_0 = 0$. In addition, without loss of generality, we normalize $a_0 = 0$. With these specifications, only $(2N + 1)$ parameters need to be specified in a county with $N$ competing MA contracts, more specifically, the parameters:

$$\{a_1, a_2, \ldots, a_N, c_1, c_2, \ldots, c_N, \gamma.\} \tag{2.14}$$

There are two possible approaches for the determination of these parameters. The first approach, employs the publicly available data for the plans' premium bids, as well as the number of enrollees in the different plans and the county's capitation rate. As we will demonstrate, these data allow us to specify a system of equations and inequalities from which the unknown model parameters in (2.14) can be backed out, with two qualifications: first, the system of equations (and inequalities) has one more unknown parameter compared to the number of equations/ inequalities. Second, it is theoretically possible that some of the plans' cost rates can only be specified within a given interval, rather than their exact value being computable. However, this theoretical complication never arose, in practice. We therefore show that all of the model parameters in (2.14) can be determined from the system of equations/inequalities by "importing" an exogenous estimate for a *single* parameter.

The *second* approach is more traditional: it calls for a precise specification of the intercept value in terms of plan-and-county attributes. Such a specification typically employs a common structure for all counties to reduce the number of parameters that need to be identified. Thereafter the parameters are estimated with an appropriate estimation method. In Sections 1.1 and 2.2, we have discussed several such estimated models.

There are advantages and disadvantages to both approaches. The (first) calibration approach has the advantage of offering relative *robustness*. Its methodology does not depend on a precise specification of the relevant plan-and-county attributes or the specific way in which they impact on the plan's utility measure. Note that all of our ultimate performance measures only depend on the *aggregate* intercept values $\{a_i\}$ rather than the underlying detailed specification of these quantities. The calibration method is, in addition, computationally much simpler than any of the above estimation methods, for example, a Generalized Method of Moments, and it generates a unique set of values rather than a combination of confidence intervals. A disadvantage, however, is that it relies on an exogenously imported parameter, while the second, traditional estimation method is fully self contained.

In view of these various advantages and disadvantages, we have pursued *both* approaches. In this section, we develop estimates based on a calibration method. In the next section, we do so based on the second method where parameter combinations are drawn from joint distributions extracted from the estimated results in Curto *et al.* [2015], arguably the most recent and comprehensive estimated model, and, to our knowledge, the only one which has adopted the precise non-linear way in which our-of-pocket expenses depend on the nominal premium bids.

Once a model has been specified, via either of the above methods, we pursue the same counterfactual studies to assess the consequences of any alternative capitation structure, all of which amount to an alternative specification of the function $g(\cdot)$, see (2.1a)–(2.1d), or an alternative choice for the response function $f(\cdot)$.

### 2.5.1 The Calibration Method

Our study is based on the county-by-county data in 2010. For this calendar year, private insurance companies submitted 14576 so-called "contracts", nationwide. This implies that, in an average county, 5.35 companies compete with each other as well as the traditional FFS plan, for the patronage of the county's beneficiaries. We have focused on all counties with 2 or more contracts; this represents 2478 out of the total of 2727 counties in the United States, covering approximately 41 million beneficiaries, with a total 14327 contracts.

Enrollment data for each plan-county combination, are publicly available, from the CMS.[4] Table 2.1 summarizes statistics for key factors in this study, the capitation rate, the (nominal) premium bids of the MA plans, cost rates for the FFS plans, and the market shares of these plans. Note that the capitation rates offered to the MA plans were, on average 10% higher than the cost rates under the traditional FFS plan. Also, while the MA plans attracted 24% of the national population, their average market share across the different counties was 15%. Thus, discrepancy is due to the fact that the MA plans are particularly prominent in relative few, but densely populated counties, see Figure 2.4 for state-by-state comparison.

Substituting $g(b) = C, \gamma_i = \gamma$ for $i = 1, 2, \ldots, N$ and $\gamma_0 = 0$ into (2.5), it follows that:

$$\log(d_i(b)) - \log(d_0(b)) = a_i - \gamma f(b_i - C), \quad i = 1, 2, \ldots, N, \quad \text{or}$$
$$a_i = \log(d_i(b)) - \log(d_0(b)) + \gamma f(b_i - C), \quad i = 1, 2, \ldots, N. \tag{2.15}$$

In addition, it was shown in Federgruen and Lu [2016a] that under an exogenous constant capitation rate, (any) equilibrium bid vector $\mathbf{b}^*$ is always an interior part of the feasible price cube $[b^{min}, b^{max}]$. Moreover, each of the profit functions, $\pi_i(\mathbf{b}), i = 1, 2, \ldots, N$, defined by (2.6), is differentiable everywhere with respect to $b_i$, with the exception of the value $b_i = C$. Finally, Proposition 1 in

[4]The enrollment data for each MA plan in each county can be obtained from the Contract/Plan/State/County enrollment data. The number of beneficiaries enrolled in the FFS plan, in each county, are contained in FFS Data 2010.

Table 2.1: Summary Statistics

| Variable | Obs. | mean | std | min | max |
|---|---|---|---|---|---|
| Capitation ($) | 2,478* | 788.682 | 70.4933 | 736.47 | 1299.79 |
| Nominal Premium ($) of individual MA plan | 14,327** | 763.098 | 68.6997 | 415.87 | 1118.875 |
| Price of MA plan as a ratio of Capitation | 14,327 | 0.9562 | 0.0995 | 0.3806 | 1.3535 |
| Cost ($) of tranditional FFS plan | 2,478 | 714.3036 | 77.9865 | 522.93 | 1286.64 |
| Standard cost of FFS plan as a ratio of Capitation rate | 2,478 | 0.9074 | 0.0817 | 0.5796 | 1.2246 |
| Market share for each MA plan | 14,327 | 0.0252 | 0.0406 | 0 | 0.4673 |
| Market share for FFS plan | 2,478 | 0.8502 | 0.1156 | 0.3378 | 0.9962 |

Note: * represents 2,478 counties, ** represents 14,327 total MA contracts in counties with two or more MA contracts.

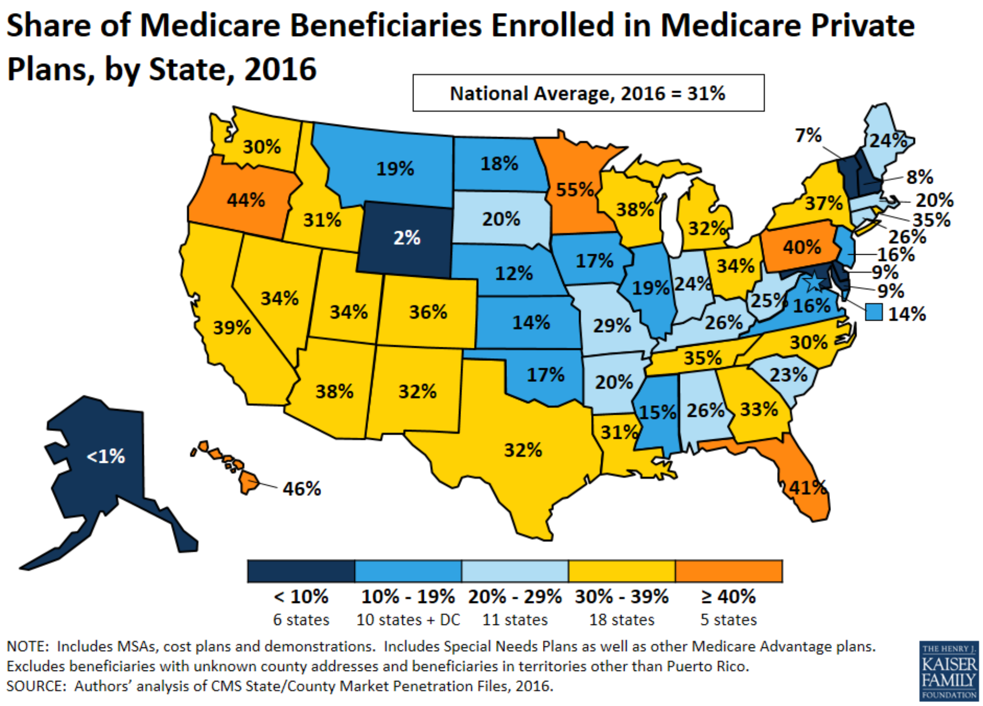



Figure 2.4: Private MA Market Shares

Federgruen and Lu [2016a] shows that the profit function $\pi_i(\cdot)$ is quasi-concave in its own price variable. These three properties imply that if an insurance firm $i$ selects a bid $b_i^* \neq C$

$$0 = \frac{\partial \log(\pi_i(\mathbf{b}^*))}{\partial b_i} = \frac{1}{b_i^* - c_i} + \frac{\partial \log(d_i(\mathbf{b}^*))}{\partial b_i} = \frac{1}{b_i^* - c_i} - \gamma f'(b_i^* - C)(1 - d_i). \tag{2.16}$$

Solving for $c_i$, we obtain for all $i = 1, 2, \ldots, N$:

$$c_i = b_i^* - \frac{1}{\gamma f'(b_i^* - C)(1 - d_i)}, \qquad \text{where } b_i^* \neq C. \tag{2.17}$$

Only when the firm decides to set its premium exactly equal to the capitation rate, i.e., $b_i^* = C$, does (2.17) fail to hold. (Indeed, the response function $f(\cdot)$ fails to be differentiable.) However, since the profit function $\pi_i(b_i, b_{-i})$ is quasi-concave in $b_i$ and achieves its maximum for $b_i = b_i^*$, we obtain that

$$\frac{\partial^- \log(\pi_i(\mathbf{b}^*))}{\partial b_i} = \frac{1}{b_i^* - c_i} + \frac{\partial^- \log(d_i(\mathbf{b}^*))}{\partial b_i} = \frac{1}{b_i^* - c_i} - \gamma\beta(1 - d_i) \geq 0, \tag{2.18}$$

$$\frac{\partial^+ \log(\pi_i(\mathbf{b}^*))}{\partial b_i} = \frac{1}{b_i^* - c_i} + \frac{\partial^+ \log(d_i(\mathbf{b}^*))}{\partial b_i} = \frac{1}{b_i^* - c_i} - \gamma\alpha(1 - d_i) \leq 0. \tag{2.19}$$

Where $\partial^-$ and $\partial^+$ denote the left-hand and right-hand derivative, respectively. Thus, when $b_i^* = C$, the plan's cost rate $c_i$ can only be determined within an interval:

$$b_i^* - \frac{1}{\gamma\beta(1-d_i)} \leq c_i \leq b_i^* - \frac{1}{\gamma\alpha(1-d_i)}. \tag{2.20}$$

Thus, when no plan selects a premium bid that is exactly equal to the capitation rate, (2.15) and (2.17) provide a system of equations from which the unknown $\{a_i, c_i, i = 1, 2, \ldots, N\}$ in (2.14) can be determined directly, assuming an estimate for the price sensitivity coefficient $\gamma$ can be obtained. (Note that the capitation rate $C$ is publicly available.) For the latter, we have imported the price sensitivity coefficient obtained in Nosal [2012], i.e., $\widehat{\gamma} = 0.013$. Similar estimates for this coefficient were obtained in other MNL models, in particular, Dowd *et al.* [2003] with $\widehat{\gamma} = 0.019$ and in Curto *et al.* [2015] with $\widehat{\gamma} = 0.014$ (See specification (4) in Table 5.)

When $b_i^* = C$, (2.20) provides an interval in which the plan's cost rate $c_i$ falls. The width of this interval is determined by the rebate percentage; the closer the rebate percentage is to 100%, the smaller the interval is, approaching a single point value when a full rebate is provided for the cost saving. We have observed that no insurance plan specified a premium that was exactly equal to the exogenously specified capitation rate.

In conclusion, the observed premia and market shares provide enough information to determine all of the parameters in the competition model, modulo a *singular* degree of freedom. As mentioned, to remove the latter we have adopted Nosal [2012]'s estimate of the price sensitivity coefficient $\gamma$. With all model parameters specified, we have computed the price equilibrium that arises when the capitation rate is set endogenously as the lowest or second lowest bid or a weighted average of the bids, similar to several reform plans discussed in Section 1.1 and Section 2.3.

While the above describes our basic model for the Medicare Advantage market, we have also investigated a variant where the market is segmented into two customer classes: The *first* segment consists of those beneficiaries who, in the prior calendar year 2009, subscribed to the traditional FFS plan. The *second* segment consists of the remaining eligible Medicare participants, i.e., those who in 2009 enrolled in a private MA plan. We applied this segmentation to insert a search or inertia cost in the plans' utility measures, whenever an individual considers switching from the FFS plan to a private MA plan or vice versa. The presence of such inertia costs has been widely observed in the marketing literature, for example Dube *et al.* [2009]; it is all the more likely to prevail in the MA market with an elderly population choosing among fairly complex alternatives. Indeed, Nosal [2012] has focused on estimating the magnitude of this inertia effect and has found it to be significant. The segmentation gives rise to a Mixed MultiNomial model (MMNL), of course with the additional complication of utility measures being dependent on *net* rather than *gross* premia via a non-linear response function $f(\cdot)$. In the Appendix, we explain, how the remaining parameters in the demand functions, as well as the marginal cost rates can be determined in this MMNL model, assuming, once again, that the observed price vector is an equilibrium under exogenously specified capitation rates. With these parameters specified, we, again, compute the equilibrium that arises when the capitation rate is determined under various endogenously determined capitation rates.

We have computed the equilibria by applying the tatônnement scheme specified in Section 2.4.1, starting from a randomly selected price vector in the feasible price space. While, as mentioned, we cannot guarantee, on theoretical grounds, that the equilibrium is unique, we have verified this numerically, by repeating the tatonnement scheme from many randomly selected starting points, observing convergence to a unique price vector, throughout.

### 2.5.2 Results

We have computed the price equilibrium in each of the 2478 counties, that arises when the capitation rate is determined as the lowest, or second lowest bid, and compared these equilibria with the premia that prevailed (in 2010) under the existing system. The equilibria were computed, both under Plan I ($\gamma_0 = 0$) and Plan II ($\gamma_0 = \gamma$). In this subsection, we report on results for Plan II only.

Table B.1 in Appendix B reports on the average price results, across all 2478 counties, of the above described equilibrium calculations. We have computed a weighted average of the counties' results, with the counties' number of eligible participants as the weight factor. The first segment of the Table displays the results in the absence of search or inertia costs; the second (third) segment displays the same, assuming this cost value equals 2 and 4, respectively. (Nosal [2012] obtained an estimate of approximately 4 for this parameter, but reports on counterfactual studies based on various values between 0 and 4.) Each segment exhibits, first of all, the actual market results in 2010, which may also be interpreted as the equilibrium under the prevailing exogenous capitation rates. The second and third column in each table segment displays the same results, when the capitation rate is specified endogenously as the lowest and second lowest premium, respectively.

In addition to displaying the average cost value of the FFS plans and the average capitation rates under the three subsidy schemes, Table B.1 exhibits the average premium of the lowest, second lowest and third lowest plans, as well as the overall average premium, again under each of the three subsidy schemes. In addition to the absolute premium values, it is also of interest to display the various premium values as a percentage of the prevailing FFS cost value. Figure 2.5 exhibits the weighted average value of these percentages, for the lowest, second lowest and third lowest bid plan, as well as the capitation rate.

Figure 2.5: Price as percentage of FFS cost (without inertia cost)

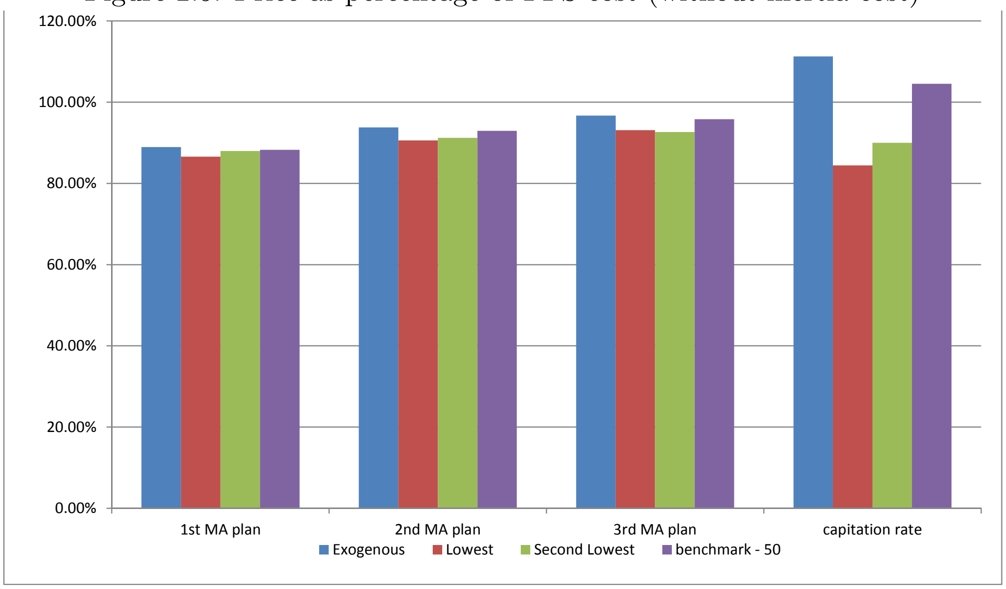
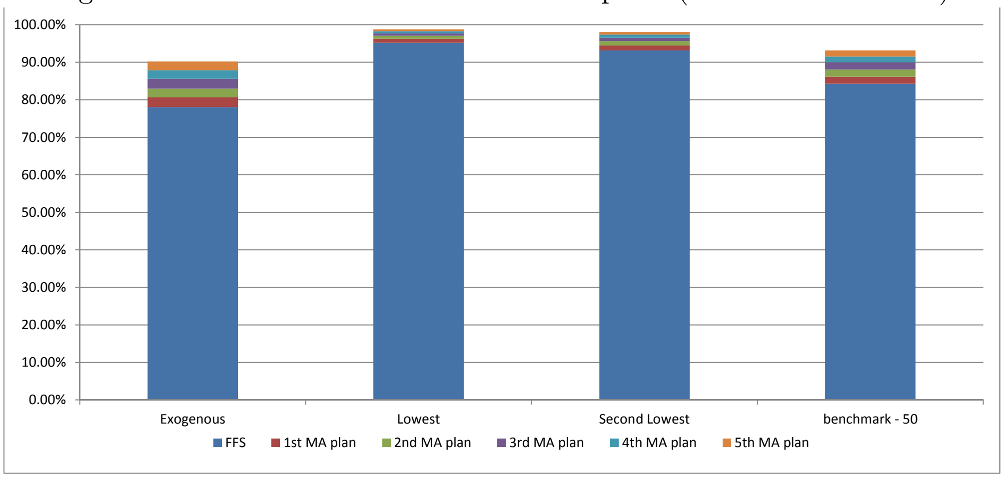



Table B.2 in Appendix B shows that the intensified competition among the private insurers, under an endogenously specified capitation rate, translates into a reduction of the market share of the traditional Medicare (FFS) plans from 78% to 73%, but only in the absence of switching costs; under an inertia cost value of 4, the average FFS market share drops by one percentage point only. The lowest priced plans increase their market share from 2.58% to 3.17%, but each individual MA plan continues to have a relatively small market share. Thus, market concentration remains low, boding well for the continuation of a healthy competitive environment. The results also indicate that beneficiaries consider many non-price related attributes in their plan choices. The same results are shown graphically in Figure 2.6.

Focusing on the case without search/inertia costs, we observe that the average capitation rate is reduced from \$838 to \$641 or \$683, depending upon whether it is specified as the lowest or second lowest bid. Since both the Wyden-Ryan, the Domenici-Rivlin, and the Rivlin-Daniel plans adopt

Figure 2.6: Market share for FFS and MA plans (without inertia cost)



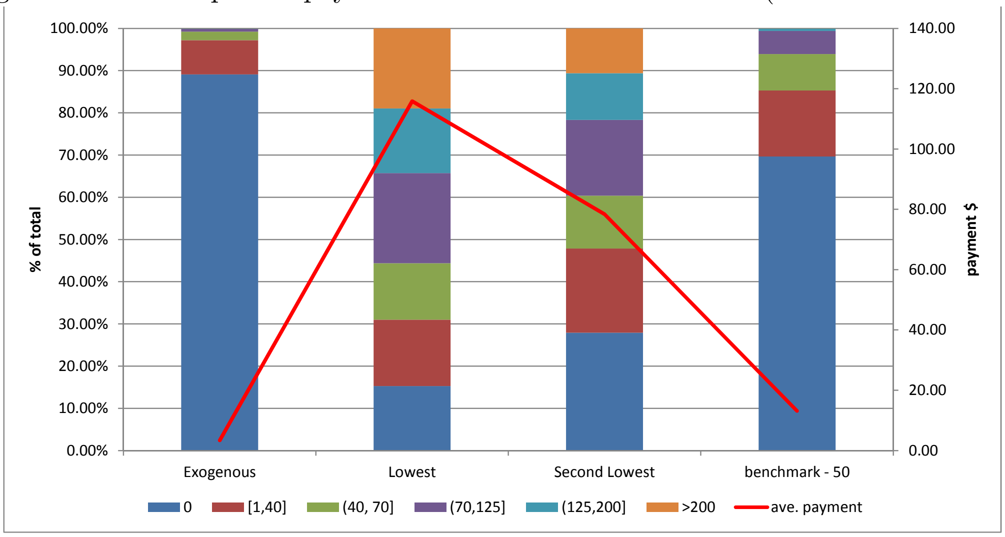

the latter scheme, we confine our summary conclusions to the latter. The reduction of the capitation rates amounts to *a cost saving for the government of no less than 16.5%.*[5] (The average capitation rate exceeded the average cost in FFS plans by close to 10%.) Under the existing capitation scheme, the average of the second lowest premia amounted to \$706, a 15.8% reduction compared with the average prevailing capitation rate. This demonstrates that the standard way of estimating the cost savings results in a significant underestimation of the savings potential: As may be expected, insurance companies react to the new subsidy scheme by bidding more aggressively and reducing their premia. Indeed, as shown in Figure 2.5, the average second lowest premium goes down from 94% to 91% of the average FFS cost value.

For those beneficiaries who choose to stay with, or newly enroll in a traditional FFS plan, the average of their out of pocket costs would amount to \$70 per month or \$840 per year, similar to earlier estimates in Song et al.(2012). Similarly, the average out of pocket costs among all Medicare beneficiaries would be \$64 per month or \$768 per year, see Table B.3 in Appendix B. These results are very similar, when incorporating an assumed search/inertia cost of 2 or 4; the prevalence of an inertia cost component increases the premia, and hence the capitation rates, slightly, as it mitigates the competitive intensity.

Finally, Table B.3 shows what percentage of all beneficiaries incurs an out of pocket expense in five specific cost buckets. Under the second lowest bid scheme, a majority would continue to pay

[5]The percentage saved is double that estimated by Feldman *et al.* [2012].

Figure 2.7: Out-of-pocket payment for FFS and MA enrollees (without inertia cost)




less than \$40 per month, while 97% do under the current exogenous subsidy scheme. Twenty-nine percent of beneficiaries would continue to pay nothing or get a rebate. Figure 2.7 displays the cumulative percentage of beneficiaries who pay less than a given amount. For example, about 50% (or 65%) of beneficiaries pay less than \$40 (or \$75) per month under the payment scheme under which the capitation rate is determined by the second-lowest bid.

## 2.6 Estimating the Impact of Medicare Reform Proposals: An Estimation Model

Our second set of counterfactual studies is based on a model contained in Curto *et al.* [2015]; specifically, model 4 in Panel A of Table 5, ibid. This is a MNL model with utility measures specified as in (2.2). In this model, the price sensitivity coefficient is estimated as $\tilde{\gamma} = 0.014$, very close to the value $\tilde{\gamma} = 0.013$ used in the calibration model in Section 2.5.1, itself "imported" from the estimation model in Nosal [2012]. In this model, the intercept coefficients $\{a_i\}$ are specified as a linear combination of the plan quality scores, indicator variables describing whether the contract covers Part D benefits, and supplemental benefits such as vision or dental coverage, as well as indicator variables for the calendar year and the contract number. See equations (11) and (12) in Curto *et al.* [2015] for more details. Curto *et al.* [2015] have estimated a nested sequence of models, with progressively more explanatory (non-price) variables added to the utility measures. Model (4) represents the richest of these five specifications, except for the omission of indicator variables for county-contract combinations.

Due to the confidentiality agreement between Curto *et al.* [2015] and the consulting firm which provided them with the data, we were unable to obtain the actual intercept and cost rate values $\{a_i, c_i\}$ for each contract in the various counties. Instead, Curto *et al.* [2015] kindly shared with us the joint distribution of the parameter pair $\{a_i, c_i\}$ and this differentiated by the number of contracts $N$ in the county, and one of 5 consecutive ranges for the capitation rate. More specifically, we focused on all counties with $N \geq 2$ contracts and divided them into 10 categories: (1)$N = 2$, (2)$N = 3$, ..., (8)$N = 9$, (9)$10 \leq N \leq 14$ and (10)$N \geq 15$. Each of these counties are further subdivided into 4 subsets corresponding with the largest, second largest, third largest, and smallest inter-quintile range of the capitation rate value $C$. For each of these 40 sets, we obtained a $5 \times 5$-table with the joint pdf of the $\{a_i, c_i\}$ parameters, discretized on a two-dimensional grid generated by the quintiles of the two parameters in the set. (Curto *et al.* 2015 also provided us with the quintiles of nominal bid values $\{b_i\}$ as benchmarks, to compare our results with.)

Based on these 40 joint distributions, we simulated for each of the 2478 counties with $N \geq 2$ contracts, a sample of 1000 market instances and computed the (unique) equilibrium for each instance. To generate an instance, for a given county, we endowed each of the competing contracts with an $\{a_i, c_i\}$ pair of parameters. These pairs were generated independently from the above joint distribution pertaining to the set corresponding to the county's known number of contracts $N$ and capitation rate $C$. More specifically, the parameter pair $\{a_i, c_i\}$ for each contract was generated by the following two step procedure: (a) in the first stage, one of the 25 above described rectangles in the two-dimensional parameter space is generated from the joint distribution, (b) a specific point within the generated rectangle is drawn from the two-dimensional uniform distribution on this rectangle.

Our results are summarized in Tables 2.2 and 2.3 (for Plan II ($\gamma_0 = \gamma$) and Plan I ($\gamma_0 = 0$) respectively) and show results to that are consistent with those derived from an calibration model (Both tables reflect scenarios without switching costs, i.e., $S = 0$, Tables B.5-B.6 in Appendix show the same results for the case where $S = 2$ and $S = 4$, under Plan I and Plan II.) We display the equilibrium results when the capitation rate is determined *endogenously* as the lowest or second lowest bid, or *exogenously* at $50 below the actual 2010 value.

Focusing on Table 2.2, the equilibrium results obtained from the estimated model shows that the capitation rates are somewhat smaller than those obtained from the calibrated models. For example, if the capitation rate is determined as the second lowest premium bid, the average capitation rate is $645 as opposed to $683. Based on the estimated model, the government's cost would be reduced by close to 21% as opposed to the 16.5% estimate we obtained from our calibrated model. The average FFS bid under this scheme is estimated to be $10, or 1.3% lower than the estimate we obtained in the calibrated model. Moreover, the average premium bid is virtually identical under the estimated vs calibrated model ($745 vs $742). Under the estimated model, the average market share of the traditional FFS program is predicted to increase by 5%, while it is predicted to *decrease* by the same in the calibrated model. In the calibrated model, the average monthly out-of-pocket cost is calculated at $64 in the estimated model, versus $103 in the calibrated model. Note that, under Plan II, all participants in the traditional FFS program pay an out-of-pocket cost, similar to these enrolled in the MA plans.

Under Plan I, see table 2.3, there is similar consistency between the results obtained from the estimated versus the calibrated models. The average premium bid is $755 in the estimated model and $745 in the calibrated model. The average capitation rate reduction is again approximately 23% and 18.5% in the two respective models. The market share percentage of the FFS program now increases to the mid 90s, because, under Plan I, with a 20% decrease in the capitation rate, there is a major difference in the net premium for MA plan enrollees versus those opting for the traditional FFS program. The average out-of-pocket cost among all beneficiaries is now $2.82 in the estimated model, and $4.27 in the calibrated model. Last but not least, government costs are estimated to decrease by 8%, irrespective of whether the estimate are based on the estimated or the calibrated models.

## 2.7 Conclusion

In this paper, we have applied a new type of price competition model to the Medicare market. The main distinguishing features of the competition model includes the fact that the utility measures associated with the competing plans depend on the *net premium*, which equals the nominal premium

Table 2.2: Estimation Result via Estimation (PlanII and no inertia cost)

| | | Estimation Model | | | Calibration Model | | |
|---|---|---|---|---|---|---|---|
| | Exogenous | Lowest | Second | Exogenous ($C-50$) | Lowest | Second | Exogenous ($C-50$) |
| **Prices** | | | | | | | |
| FFS | 752.96 | 736.61 | 736.96 | 741.61 | 746.92 | 747.36 | 752.43 |
| Capitation | 837.68 | 603.26 | 645.43 | 787.68 | 640.64 | 683.32 | 787.68 |
| ave. price | 753.82 | 744.56 | 744.88 | 748.76 | 740.95 | 742.47 | 752.54 |
| **Market Share** | | | | | | | |
| FFS | 78.05% | 82.64% | 83.22% | 87.51% | 72.88% | 73.25% | 77.51% |
| 1st MA plan | 2.58% | 1.42% | 1.35% | 1.02% | 3.77% | 3.17% | 2.63% |
| 2nd MA plan | 2.34% | 1.42% | 1.37% | 1.01% | 3.23% | 3.19% | 2.45% |
| 3rd MA plan | 2.63% | 1.41% | 1.35% | 1.02% | 3.20% | 3.25% | 2.69% |
| **Out-of-pocket payment per month** | | | | | | | |
| 0 | 89.12% | 1.52% | 11.43% | 70.48% | 16.49% | 28.71% | 70.13% |
| (0,40] | 8.07% | 12.58% | 18.31% | 17.15% | 16.35% | 21.48% | 15.37% |
| (40, 70] | 2.03% | 12.05% | 14.94% | 8.16% | 14.61% | 14.68% | 8.40% |
| (70,125] | 0.68% | 24.04% | 20.01% | 3.84% | 23.50% | 19.12% | 5.39% |
| (125,200] | 0.07% | 22.68% | 17.16% | 0.36% | 14.91% | 9.65% | 0.67% |
| $> 200$ | 0.03% | 27.14% | 18.15% | 0 | 14.13% | 6.36% | 0.04% |
| ave. payment | 3.37 | 141.30 | 102.75 | 11.24 | 100.31 | 64.25 | 13.06 |
| **Government payment ($) per month** | | | | | | | |
| payment | 815.87 | 599.96 | 641.82 | 775.75 | 641.80 | 683.50 | 776.12 |
| % of Exogenous | | 73.54% | 78.67% | 95.08% | 78.66% | 83.78% | 95.13% |

Table 2.3: Estimation Result via Estimation (PlanI and no inertia cost)

| | | Estimation Model | | | Calibration Model | | |
|---|---|---|---|---|---|---|---|
| | Exogenous | Lowest | Second | Exogenous ($C-50$) | Lowest | Second | Exogenous ($C-50$) |
| **Prices** | | | | | | | |
| FFS | 752.96 | 754.62 | 754.47 | 742.88 | 756.32 | 756.88 | 753.47 |
| Capitation | 837.68 | 599.73 | 642.47 | 787.68 | 638.46 | 681.14 | 787.68 |
| ave. price | 753.82 | 754.77 | 754.82 | 751.25 | 742.84 | 745.43 | 755.84 |
| **Market Share** | | | | | | | |
| FFS | 78.05% | 97.73% | 96.54% | 85.28% | 95.20% | 93.15% | 84.27% |
| 1st MA plan | 2.58% | 0.23% | 0.36% | 1.20% | 1.03% | 1.31% | 1.87% |
| 2nd MA plan | 2.34% | 0.23% | 0.36% | 1.21% | 0.82% | 1.16% | 1.88% |
| 3rd MA plan | 2.63% | 0.23% | 0.35% | 1.20% | 0.64% | 0.88% | 1.97% |
| **Out-of-pocket payment per month** | | | | | | | |
| 0 | 97.38% | 95.98% | 95.38% | 97.35% | 97.73% | 96.55% | 96.33% |
| (0,40] | 1.82% | 1.26% | 1.84% | 1.13% | 0.04% | 0.30% | 2.12% |
| (40, 70] | 0.45% | 0.75% | 1.09% | 0.84% | 0.12% | 0.63% | 1.53% |
| (70,125] | 0.26% | 1.37% | 1.24% | 0.51% | 0.83% | 0.85% | 0.01% |
| (125,200] | 0.06% | 0.57% | 0.41% | 0.15% | 0.40% | 1.02% | 0.00% |
| $> 200$ | 0.03% | 0.07% | 0.04% | 0.02% | 0.89% | 0.65% | 0.00% |
| ave. payment | 0.92 | 2.96 | 2.82 | 1.38 | 3.73 | 4.27 | 1.13 |
| **Government payment ($) per month** | | | | | | | |
| payment | 815.87 | 751.05 | 750.55 | 758.40 | 751.36 | 751.82 | 759.61 |
| % of Exogenous | | 92.05% | 91.99% | 92.96% | 92.09% | 92.15% | 93.10% |

minus the government's subsidy or capitation rate; moreover, utilities depend in a *non-linear* way on the net premium. In the current system, the capitation rates are exogenous specified and pre-announced; in various reform plans, they would be replaced by an endogenously subsidy, determined as the lowest or second lowest bid, or a weighted average of all bids. Finally, the utility measures may include a "switching cost" incurred when a beneficiary switches between a traditional FFS plan and a Medicare Advantage plan, or vice versa.

We have applied these models to the 2010 county-by-county data. We employed two different methodologies to derive the parameters in the various competitive models: one method employs a calibration technique and the other parameter distributions that were estimated by Curto *et al.* [2015]. Computing the price equilibria in the various models under various subsidy structures, we have been able to estimate the impacts on price equilibria, capitation rates, market shares, government expenditures, and out-of-pocket costs. Our results exhibit major government cost savings under several of these reform plans, with modest increases in out-of-pocket costs.

## Acknowledgement

We are greatly indebted to Jon Levin of the Stanford Business School for providing us with the parameter distributions which are the basis of the counterfactual studies in Section 2.6. We also acknowledge the major help Evan Mast of the Stanford Business School provided in generating the required data and distributions from Curto *et al.* [2015]'s data base.

# Chapter 3

# Framework Agreements in Procurement: An Auction Model and Design Recommendations

Yonatan Gur, Lijian Lu, and Gabriel Y. Weintraub

## 3.1 Introduction

### 3.1.1 Background and motivation

Governments around the world spend billions of dollars every year buying a wide range of products and services from private firms. While standard auctions are often used to allocate contracts in these procurement processes, recent years have seen a tremendous increase in the adoption of an alternative class of mechanisms in various public procurement settings: the so-called *framework agreements (FAs)*, also called indefinite-delivery/indefinite-quantity (IDIQ) contracts in the United States. FAs award tens of billions of dollars worth of contracts annually around the globe and constitute a steadily increasing fraction of governments' procurement processes. For example, FAs awarded €85 billion in 2010 in the European Union only, accounting for 17% of the total value of all contracts awarded, and their use has increased in the EU at an average rate of 18% since 2006.[1]

Broadly speaking, FAs are anticipated arrangements for the delivery of goods and services over a certain period of time.[2] Motivated by the increasing adoption of this class of mechanisms in practice, we aim at providing a better understanding of FAs by introducing a novel auction model to study a salient feature of these mechanisms, and proposing concrete design recommendations to improve their performance. This work is the result of a collaboration with the Chilean government procurement agency *Dirección ChileCompra* (ChileCompra for short) that buys around 10 billion dollars worth of products and services every year, of which 2 billion are bought using FAs.

**What are FAs?** Consider a government that is interested in buying computers for its public agencies (e.g., schools, hospitals) for the next two years, a time period during which many demand requests can be expected. On one hand, running an auction whenever such a request arises is administratively expensive, considering that requests may be frequent and of a small volume. On the other hand, letting each agency run its own procurement process does not exploit the central government's bargaining power and buying know-how. FAs strike a balance between a decentralized

[1] Data from "Public Procurement in Europe," http://ec.europa.eu/internal_market/publicprocurement/docs/modernising_rules/cost-effectiveness_en.pdf.

[2] The European Parliament defined FAs as "an agreement between one or more contracting authorities and one or more economic operators, the purpose of which is to establish the terms governing contracts to be awarded during a given period, in particular with regard to price and, where appropriate, the quantity envisaged". (The Directive 2004/18/EC of the European Parliament and of the council of March 31, 2004.)

procurement process and the central government's bargaining power. Suppliers, on their end, are motivated to participate in FAs due to the large demand that might be associated with them.

A typical FA is composed of two stages: in the *auction stage*, an auction-type mechanism, typically first-price, takes place to select one supplier as the *FA winner* for a given product or service. The FA winner is required to sell over the time horizon of the FA at the bid price determined at this first stage. Then, in the *buying stage*, various government agencies may buy the product from the FA winner as needs arise. It is common that government agencies have the obligation to buy from the FA winner unless they can provide evidence of a more convenient procurement option in the open market (that we also refer to as the outside or the spot market). Overall, an FA can be viewed as a *government call option* to buy at a predetermined price over the time horizon.

In this paper we focus on a distinctive feature that FAs exhibit relative to running a standard first-price auction whenever a need arises. FA bidders face significant cost uncertainty; while the price of a product or service is locked at the auction stage, the suppliers' costs may change over the buying stage. Anecdotal and empirical evidence suggests that providers "charge" for this uncertainty through higher bids. Guillermo Burr, head of Research at ChileCompra, says "we wanted to better understand why in some categories standard auctions resulted in lower prices relative to FAs, and how to alleviate this problem and reduce ChileCompra's buying prices."

### 3.1.2 Main contributions

Despite their practical importance, there is little academic research on FAs. In the current paper we develop a model for FAs that considers the *cost uncertainty* faced by suppliers and the resulting bidding incentives. Our paper contributes to the literature on procurement mechanisms in the interface between operations and economics, and at the same time has concrete practical implications. The contributions of the paper can be categorized along the following three dimensions.

**(i) Modeling and bidding incentives.** We introduce an auction model for FAs that generalizes standard auction models to incorporate cost uncertainties faced by suppliers at the auction stage. In our model all suppliers face a common cost that is unknown at the auction stage and is realized at the buying stage (for instance, the cost of gas in a transportation service). Then, a buyer has the option to buy from the FA winner at the agreed price, or to buy from the outside market that exhibits a similar structure to the bidders' costs; specifically, it also incorporates the common cost component, as outside providers may also need to incurr this cost to provide the good. Given this structure, we identify that FAs are subject to a sort of *winner's curse* (cf. Krishna 2002), because the events under which the FA winner sells the product are positively correlated with high cost realizations. Intuitively, the FA option is exercised when the locked-in price is attractive relative to the spot market price; this coincides with large cost realizations for suppliers. The manner bidders react to the *FA curse* in equilibrium formalizes practitioners' intuition that FAs may result in larger prices relative to running standard first-price auctions as needs arise. With this motivation, we introduce various variants of FAs that alleviate the strategic response to the FA curse, reducing expected payments to suppliers.

**(ii) Analysis.** We study the Bayes Nash equilibrium (BNE) of the game of incomplete information between sellers induced by different FA variants. We compare the expected buying prices among these FA variants (as well as first price auctions) using an envelope theorem approach. By itself, the latter analysis has novelties relative to standard mechanism design, as the outside option given by the spot market price is endogenous and depends on suppliers' private information, resulting in technical challenges when applying standard approaches. In addition, the analysis of the *restricted-flexible FA* discussed below is novel as it requires solving a dynamic programming problem embedded in an auction model. We complement our theoretical results by supporting numerical experiments that demonstrate the robustness of our findings.

**(iii) Practical Design Prescriptions.** A significant contribution that particularly distinguishes this paper from previous related work in auctions is a series of practical suggestions for the design of FAs that arise from our results and that we summarize below.

We show that monitoring the price offered by the FA winner in the outside market and forcing the FA winner to match it whenever it is lower than the winning bid, significantly reduces expected buying prices. Hence, governments may capture significant value by centrally monitoring prices charged by FA winners in the open market. Considering that thousands of products are bought through FAs, central monitoring operations might be associated with practical challenges that may prevent their adoption by procurement agencies despite the price reduction associated with it. As an alternative, we show that using a *perfect price index* in which the auctioneer perfectly observes the realization of the common cost and indexes the bid of the FA winner to its changes lowers expected buying prices in many settings of practical interest. Hence, if possible, governments should make an effort to invest in finding and implementing price indexes for the random common part of suppliers' costs.

Price indexes are typically available in practice when the common cost is associated with a commodity (such as gas). However, for many of the goods and services procured through FAs (e.g., computers, office equipment) such indexes do not exist or are hard to build. Thus motivated, we study a new practical variant of the standard FA, the *flexible FA*. Here, if the FA winner has a bid larger than the spot market price offered by a different supplier, the FA winner is allowed to match it and sell the product. While it does not require central monitoring operations nor price indexes, we show that the flexible FA achieves expected buying prices that are often comparable with ones obtained by using monitoring operations, perfect price indexes, or even running first-price auctions as needs arise. Motivated by practical considerations, we also study the performance of the *restricted-flexible FA*, in which once the FA winner reduces its price to match the spot market price, he cannot increase it back again. We show that this restriction, that may be appealing for buying agencies because of the transparency it induces, does not significantly hurt expected buying prices in many scenarios of interest.

These prescriptions are already being applied by ChileCompra to improve the design of their FAs. Mr. Burr says, "These results have provided important insights regarding the design of our FAs. They have encourage our FA department to make larger efforts to build adequate price indexes. They have also motivated a larger effort in policing spot market prices FA winners offer, by allowing buyers to report low spot market prices when they observe them. They have also showed us the types of pricing flexibility that we should encourage."

### 3.1.3 Related literature

Our study relates to several papers in the literature on auction theory. As the spot market in our model plays a similar role to a random reserve price, our analysis shares some similarities with Elyakime *et al.* [1994] that compares seller's expected revenue when using a secret random reserve price relative to the optimal public reserve price. In their model, however, there is no correlation between bidders' costs and the reserve price, like in our case through the common cost component. In this sense, our study is more closely related to classic work in common value auctions and the associated 'winner's curse.' Specifically, our FA model is similar to auction models with both private and common values (see, e.g., Goeree and Offerman 2003). Apart from several technical differences between our work and the papers mentioned above, a key distinction is the operational and practical contribution that we provide for FAs, particularly those related to monitored and flexible FAs. In contrast, none of the aforementioned papers focus on FAs nor provide concrete design prescriptions for them. An additional operational aspect that is emphasized in the current paper is the impact of restricted price flexibility, which is approached by solving a dynamic programming problem embedded in an auction model.

Our study relates to a growing stream of work in operations that studies supply chains and procurement processes under different forms of uncertainty. Several studies have focused on demand uncertainty that is faced by buyers and/or suppliers: Chen [2007] and Duenyas *et al.* [2013] study optimal procurement mechanisms in a newsvendor-like setting where a buyer facing uncertain demand determines both the quantity and purchasing price through interactions with suppliers; Li and Huh [2011] consider a buyer facing uncertain demand and suppliers that need to invest in capacity before the demand uncertainty is resolved, and study whether the buyer should offer a pull or push contract; Zhang [2010] also studies a procurement mechanism in a supply chain setting, but includes supplier delivery performance and price-sensitive market demand.

In addition, Schummer and Vohra [2003] study the mechanism design problem of a buyer that can procure purchase options from capacity constrained sellers to satisfy an unknown future demand. The focus of their work is to study how options can be used to hedge against random demand when suppliers have a cost associated to reserving capacity. Instead, in our case FAs can be viewed as government call options to lock-in a price when suppliers' costs fluctuate. More broadly, while these operations papers relate to demand uncertainty, the focus of our work is on studying the impact of bidders' cost uncertainty. In that respect, our work also relates to the one of Elmaghraby and Oh [2013] that study, in a very different setting, how to structure two sequential auctions in the presence of learning-by-doing, and whether the buyer is better-off by limiting competition and contracting with a single supplier in the hope of extracting a better future price. These questions are at some level related to designing the structure of competition with the spot market price in the buying stage of FAs (e.g., whether to allow the flexibility to match it or not). In another related paper, Tunca and Zenios [2006] provide conditions under which procurement auctions are preferred over relational contracts, as similarly to FAs they identify the most cost efficient supplier.

As mentioned, the literature that directly studies FAs is limited. Subsequent to the first version of this paper, Saban and Weintraub [2015] studied another distinctive feature of FAs, namely how to optimize the trade-off between product variety and price competition. Their auction and mechanism design analysis considers an FA model with multiple imperfect substitute products, but ignores suppliers' cost uncertainties.[3] In this sense, the two papers are complementary as they study different aspects of FAs. In this paper we abstract away from key features studied in Saban and Weintraub [2015], and from other complexities that may arise in FAs (e.g., multiple products, multiple winners, non-linear costs) in order to focus on cost uncertainties faced by suppliers, their negative impact, and practical ways of eliminating this impact. However, we believe our design prescriptions are also valid in these more general settings. We note that we presented preliminary results of this work in a conference paper that appeared in a practitioners' outlet [Anonymous, 2012]. However, that paper studies a simpler model and does not contain the theoretical results that support the managerial insights and design prescriptions of the current work.

**Structure of paper.** The rest of the paper is structured as follows. In §2 we present an auction-based model for FAs. In §3 we develop BNE bidding functions for important cases of interest. In §4 we compare the expected buying prices among various mechanisms. In §5 we introduce flexible FAs and study their performance. In §6 we provide numerical analysis that complements the theoretical results of the previous sections. In §7 we conclude with key design prescriptions. Selected proofs and material appear in Appendices A and B. Additional proofs and auxiliary results are deferred to a series of Appendices that appear in an online companion.

## 3.2 The Framework Agreement Model

We model a framework agreement (FA) as a game of incomplete information between suppliers, similarly to the classical modeling approach in auction theory (see Krishna 2002, Milgrom 2004), and use pure strategy Bayes-Nash equilibrium (BNE) as solution concept (see, e.g., Mas-Colell *et al.* 1995). Consider a buyer that is interested in procuring a product/service to satisfy demand over an horizon of $T$ time periods. We assume that demand quantities are deterministic and normalized to one in each time period. Under the assumptions of our model with risk-neutral agents, all of our

[3] Albano and Sparro [2008] studies a similar trade-off to Saban and Weintraub [2015] under complete information.

results are also valid for a sequence of i.i.d. demand quantities. To simplify notation we assume that agents do not discount future payments, but our analysis can be extended to accommodate discounting.

**Suppliers and cost structure.** We denote by $\mathcal{M}$ the set of $M$ risk-neutral potential suppliers that could provide the good or service being procured. The set of potential suppliers is assumed to be fixed throughout the FA horizon. The constant marginal cost faced by bidder $i \in \mathcal{M}$ when providing the good at time period $t \in \{1, \ldots, T\}$ is given by the sum of two components: $c_i + X_t$. The first component, $c_i$, is bidder $i$'s private cost, known only to himself at $t = 0$. The private costs $\{c_i : \ i \in \mathcal{M}\}$ are assumed to be i.i.d. random variables, each with distribution function $F$, continuous density function $f$, and finite, non-negative support $[\underline{c}, \bar{c}]$. The second cost component, $X_t$, is common to all bidders and its realization for all $t \geq 1$ is unknown at $t = 0$. The marginal distribution function of $X_t$ is denoted by $F_{X_t}$, with continuous density function $f_{X_t}$, and a finite, non-negative support $[\underline{x}, \bar{x}]$, for all $t$. The stochastic process $\mathcal{X} = \{X_t : t \geq 1\}$ is independent of the private costs $\{c_i : \ i \in \mathcal{M}\}$.

The private cost $c_i$ represents idiosyncratic characteristics of supplier $i$, such as managerial ability, logistics and production costs, which for the most part do not change over time. On the other hand, the cost $X_t$ is common, and interpreted as being related to the price of inputs that all suppliers require in order to provide the product or service. The process $\mathcal{X}$ is random as these prices may change over time. To illustrate this cost structure, consider the provision of a transportation service at period $t \geq 1$. Costs associated with the logistics of the firm and its transportation network are private and assumed to be fixed over time; these costs are represented by $c_i$.[4] On the other hand, the costs of some inputs such as gas are common to all firms and subject to random fluctuations between $t = 0$ and time period $t$; these costs are represented by $X_t$.

**Auction stage.** At time $t = 0$, a subset $\mathcal{N} \subseteq \mathcal{M}$ of $N$ suppliers participate in the auction stage and simultaneously submit sealed bids. The lowest bid wins, and we refer to the supplier with the lowest bid as the *FA winner*. We do not model entry decisions explicitly; for simplicity, we take the auctions participants as exogenously given and we remain agnostic about this entry process. For example,

[4] This may be a particularly reasonable assumption for relatively mature offerings, such as transportation services or office supplies. While in reality, a fraction of the private costs may also be subject to uncertainty between $t = 0$ and some period $t$, we abstract away from this effect, focusing on the impact of the common cost uncertainty.

it may be that among the $M$ suppliers, the $N$ that participate in this stage have the lowest private cost realizations; in such case the FA attracts the most cost efficient providers. For example, in the context of ChileCompra, for a given product category such as food, few suppliers that recurrently participate in FAs are typically believed to be cost efficient in providing the products or services required by the FA relative to other suppliers that operate in the open market.

**Buying stage.** At any time period $t \geq 1$, after the realization of the common cost $X_t$, the buyer can buy from the FA winner at a price equal to the winning bid, or can buy from the *spot or open* market as an outside option. We assume that at every $t \geq 1$ the buyer buys from the open market if and only if the realized price of this option is lower than the winning bid.

**Spot market prices.** The bid of the FA winner is compared against an outside option, which is driven by the prices that are charged in the spot market. We assume this market consists of the aforementioned set $\mathcal{M}$ of potential suppliers, and that the terms of the FA mechanism do not affect prices in the open market.[5] In what follows we describe how spot market prices are formed, and then describe different outside options that may be considered as a function of these prices.

We assume that the price supplier $i \in \mathcal{M}$ sets in the *spot market* at time $t \geq 1$ is given by $c_i + X_t + Z_{i,t}$, where $Z_{i,t}$ is an additive markup charged by firm $i$ at time $t$. Typically, these markups are set as an equilibrium outcome resulting from competition among suppliers in the open market. The type of additive markups we assume arise, for example, from a price competition equilibrium in a classic model of monopolistic competition with a logit demand system (see Besanko *et al.* 1990 for details). In this case, $Z_{i,t} = Z_t$ for all $i$, where $Z_t$ can be interpreted as a diversity parameter; the greater its value, the more weight consumers assign to idiosyncratic taste factors in the logit demand system. While more complex competition models can be considered for the spot market, we believe this additive model with firm-specific markups $Z_{i,t}$ provides basic richness without over complicating the analysis of bidding incentives and the FA auction outcome, which is our main focus.

[5] This assumption is reasonable for most products bought by ChileCompra based on the observation that the volume of sales transacted through FAs is typically small relative to the overall volume transacted in the open national market. While there are exceptions (for example, in Chile the majority of the volume associated with certain health services such as dialysis is transacted through FAs), in most of FA categories (e.g., electronic devices, furniture, food, office supplies) the volume transacted through FAs is much smaller than the overall national transacted volume. In these cases, it may not be realistic to assume that suppliers will strategize in the spot market because of the FA.

We assume that for each $i \in \mathcal{M}$ the markups $\{Z_{i,t};\ t \geq 1\}$ are unknown at time period $t = 0$ and that $Z_{i,t}$ gets realized at time $t$. To simplify the analysis we assume that the stochastic processes $\{\boldsymbol{Z}_t = (Z_{1,t}, \ldots, Z_{M,t});\ t \geq 1\}$ and $\{X_t;\ t \geq 1\}$ are independent. (We use boldfaces to denote vectors throughout the paper.) However, our model does not rule out time correlation on these processes. For each $t \in \{1, \ldots, T\}$, we also allow correlations among the random variables $\{Z_{i,t};\ i \in \mathcal{M}\}$, but assume they share the same marginal distribution function $F_{Z_t}$ with continuous density $f_{Z_t}$, and finite, non-negative support $[\underline{z}, \overline{z}]$.

| | Naïve FA | Monitored FA |
|---|---|---|
| Concentrated market (small M) | Outside option selected randomly among all M suppliers; Probability to be sampled as the outside option is large. | The FA winner is always the one selected as the outside option. |
| Defused market (large M) | Outside option selected randomly among all M suppliers; Bidders neglect probability to be sampled as the outside option. | |

Figure 3.1: **Cases of interest for spot market price.**

**Outside Option.** Given the prices at the open market, there are different ways of determining the outside option against which the bid of the FA winner is compared. We give special attention to the following cases of interest, that are summarized in Figure 3.1. On one extreme is the *monitored FA*, in which the outside option is always determined by the price charged in the open market by the *FA winner*. Monitoring the spot market price of the FA winner is an approach taken by buying agencies that play a pro-active role not only in the initial tendering stage at $t = 0$, but also in providing outside options at the buying stage. For example, the Korean procurement agency has embraced this approach and is continuously monitoring FA winners' prices in the open market [Kang, 2013]. An attractive feature of the monitored FA is that it guarantees that the procurement agency never buys at a price that is larger than the FA winner spot market price. Note that if both bids and outside prices are ordered according to costs (i.e., more efficient suppliers are more competitive in the auction and in the spot market), then the FA winner exhibits the lowest expected spot market price among all FA participants. Moreover, if the most cost efficient suppliers participate in the FA, then the FA winner charges the lowest average spot market price among all suppliers.

Due to the high costs associated with monitoring prices of numerous goods that are being bought through FAs, ChileCompra and other agencies have not pro-actively monitored spot market prices of FA winners. In this case, while open market prices are not monitored in a *centralized* manner, the individual agencies themselves may still compare FA prices with prices they find in the open market. ChileCompra, for example, allows agencies to purchase from suppliers that did not win the FA if their price in the open market is lower than the FA winning bid. Thus, market prices may still be monitored in a *decentralized* manner by the agencies themselves. However, it is reasonable to assume (and also commonly observed in practice) that due to idiosyncratic reasons such decentralized monitoring may not use the FA winner spot market price as point of comparison, or more generally, may not identify the lowest price in the open market. For example, individual agencies may be more sensitive to their own search costs than to reducing public spending, and therefore might compare the FA price only with a single or few 'local' suppliers that are selected based on geographical proximity or convenience.

A simple model that captures such decentralized monitoring is a *naive FA*, in which the spot market price is sampled among the set of potential suppliers *randomly.* More precisely, we assume that each potential supplier $i \in \mathcal{M}$ is sampled independently from any other random quantity with probability $q_{i,M}$ and the outside option is given by the spot market price of the sampled supplier. It is worth highlighting that all the asymptotic results that compare the performance of naive FAs in diffuse markets with other types of FAs and with first price auctions presented in §3.4 and §3.5 are valid for more general sampling mechanisms in the spot market. For example, one valid alternative would be to sample a finite set of potential suppliers from the open market and set the price of the outside option equal to the lowest sampled price. To simplify notation and exposition we prove our results with the base model in which only one firm is randomly sampled at the spot market, and in §3.6 we present numerical experiments showing comparative statics regarding different sampling mechanisms and levels of prices in the spot market.

When considering the naive FA it is worthwhile to distinguish between two different kinds of markets: (i) *concentrated markets*, in which $M$ is relatively small and each bidder has a relatively large probability of being sampled as the outside option at the buying stage; and (ii) *diffused markets*, in which $M$ is large and each bidder has a relatively small probability of being sampled as the outside option at the buying stage. Formally, in diffuse markets, we assume $q_{i,M} \to 0$ as $M \to \infty$ for any $i \in \mathcal{M}$, and $\sum_{i=1}^{M} q_{i,M} = 1$, for all $M$. The histogram in Figure 3.2 depicts a distribution of the number of suppliers competing in FAs in Chile (the number of FA participants may be seen as a lower bound for the number of potential suppliers). Many FA markets in Chile include numerous suppliers (e.g., food, computers) while other markets are concentrated (e.g., airlines, dialysis services).

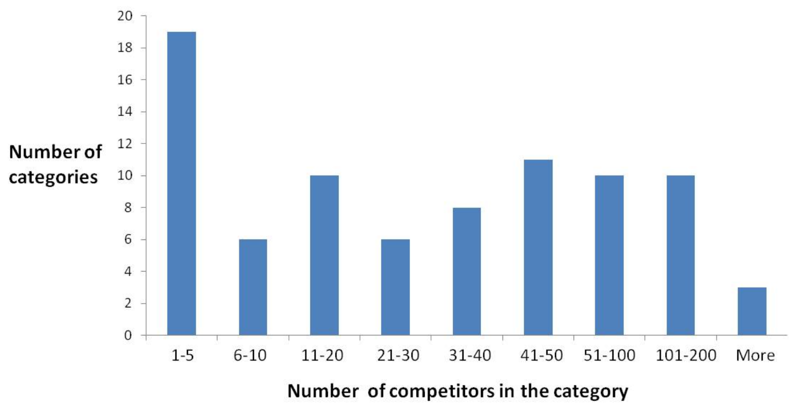



Figure 3.2: The histogram summarizes the number of competitors in 83 different product categories of FAs taken place between 2007 and 2011 in Chile. We note that in some categories it may be the case that some suppliers bid only for a subset of products/services in the category. *Source: Dirección ChileCompra.*

In the following sections we analyze the above cases of interest and show that these distinctions significantly impact both bidding behavior as well as expected payments for the procurement agency.

## 3.3 BNE Bidding Strategies

In §2 we distinguished between naive and monitored FAs, where a key difference between the two lies in the construction of the outside option against which the FA winning bid is compared at the buying stage. In the case of naive FAs we also distinguished between diffused and concentrated markets, where the difference between the two lies in the extent of competition in the open market. As we mentioned, all these regimes are of practical relevance. In this section we characterize BNE bidding functions for these cases of interest, and discuss their main elements to provide intuition on suppliers' bidding incentives in FAs. To simplify expressions in a manner that clarifies intuition we focus here on the case $T = 1$; in §4 we will return to consider the general case $T \geq 1$, when analyzing expected payments under various FAs. In this section we study symmetric, continuous, and strictly increasing BNE strategies. The bidding strategy of supplier $i$ is a mapping $\beta_i : [\underline{c}, \bar{c}] \to \mathcal{A}$, where $\mathcal{A}$ is an interval of $\Re_+$. We denote by $\boldsymbol{\beta}_{-i} = \{\beta_j, j \neq i\}$ as the vector of bidding strategies of $i$'s competitors. Throughout this section we assume the existence of such equilibria; BNE existence is concisely discussed at the end of the section (for an extended discussion see Appendix C.7).

### 3.3.1 Naive FA

In the naive FA to construct the outside option in the buying stage, a supplier is randomly sampled from the $M$ potential suppliers in the spot market. The spot market for supplier $i$ at time $t = 1$ is given by $c_i + X + Z_i$, where we ignored the time index because we are considering $T = 1$.

**Diffused Markets.** In a diffused market the number of suppliers that compete in the open market from which an outside option is drawn, $M$, is very large, while the number of firms participating in the FA, $N$, remains finite (and potentially small). Hence, the chance that a given FA participant is sampled at the spot market stage becomes negligible. We denote by $c_{(1:N),-i}$ the lowest order statistic among the costs of the suppliers that compete with $i$ in the auction stage. Given a strictly increasing competitors' strategy $\beta$, the expected profit of seller $i$ with private cost $c_i$ and bid price $b_i$ in the naive FA with diffused markets is given by:

$$\begin{aligned}\pi_i^{NDFA}(b_i, c_i, \beta) &= \lim_{M\to\infty} \mathbb{E}\left[\mathbb{I}\{b_i \le \beta(c_{(1:N),-i})\} \cdot \left(\sum_{j=1}^{M} q_{j,M}\mathbb{I}\{b_i \le c_j + X + Z_j\}\right) \cdot (b_i - c_i - X)\right] \\ &= \bar{F}^{N-1}(\beta^{-1}(b_i))\mathbb{E}\left[\mathbb{I}\{b_i \le c_j + X + Z_j\} \cdot (b_i - c_i - X)\right],\end{aligned}$$

for any $i \in \mathcal{N}$, where $\bar{F}(x) = 1 - F(x)$, and $\mathbb{I}\{\cdot\}$ denotes the indicator function. Recall that the bidder sells the good if he defeats his competitors at the auction stage and also the randomly sampled spot market price at the buying stage. Because $M \to \infty$, the bidder ignores the event in which himself is sampled in the spot market, and the cost of the sampled supplier in the spot market ($c_j$) becomes independent from the costs of the FA participants. Note that at the auction stage ($t = 0$), $X$ and $\boldsymbol{Z}$ are random and $Z_j$'s share the same marginal distributions.

To study the equilibrium strategies we could solve the ordinary differential equation (ODE) derived from the first-order condition associated to the maximization of the profit function above. However, differently to standard first price auctions, to the best of our knowledge, this ODE does not have a closed-form solution, because of the presence of the random spot market price.[6] Instead, we derive the integral equation that describes the BNE strategy, by using the envelope theorem

[6] Elyakime *et al.* [1994] provide a related discussion when studying random secret reserve prices in a first-price auction, deriving a somewhat similar integral equation to the one below.

(see, e.g., chapter 4 in Milgrom 2004 for a treatment of this approach). This equation provides more intuition regarding the equilibrium relative to the ODE. We have the following result.

**Proposition 7 (BNE bids in naive FA with diffused markets)** *Suppose that $T = 1$. Let $\beta^{NDFA}(\cdot)$ be a strictly increasing, continuous, and symmetric BNE strategy profile for the naive FA in diffused markets. Then, $\beta^{NDFA}(\cdot)$ satisfies the following integral equation for all $c_i \in [\underline{c}, \bar{c}]$:*

$$\beta^{NDFA}(c_i) \quad = \quad \underbrace{c_i + \frac{\mathbb{E}\left[X\mathbb{I}\left\{\beta^{NDFA}(c_i) \leq c_j + X + Z_j\right\}\right]}{\mathbb{P}\left[\beta^{NDFA}(c_i) \leq c_j + X + Z_j\right]}}_{(A)} + \underbrace{\frac{\int_{c_i}^{\bar{c}} \bar{F}^{N-1}(c) \cdot \mathbb{P}\left\{\beta^{NDFA}(c) \leq c_j + X + Z_j\right\} dc}{\bar{F}^{N-1}(c_i) \cdot \mathbb{P}\left\{\beta^{NDFA}(c_i) \leq c_j + X + Z_j\right\}}}_{(B)}.$$

We refer to the first two components (A) of the right-hand-side of the equation above as the *implied cost*: the expected cost of the bidder, conditional on offering a better price than the spot market. We refer to the third term (B) as the *markup* the bidder charges on top of the implied cost for having a private cost lower than $\bar{c}$. The last term is similar to the "information rent" term in a standard first price auction; however, in our model the bidder needs to defeat not only all the other bidders but also the spot market. Considering term (A), one may obtain:

$$\frac{\mathbb{E}\left[X\mathbb{I}\{\beta^{\text{NDFA}}(c_i) \leq c_j + X + Z_j\}\right]}{\mathbb{P}\left[\beta^{\text{NDFA}}(c_i) \leq c_j + X + Z_j\right]} \;=\; \mathbb{E}\left[X \middle| \beta^{\text{NDFA}}(c_i) \leq c_j + X + Z_j\right] \;\geq\; \mathbb{E}\left[X\right], \tag{3.1}$$

for any realized private cost $c_i$.[7] This expression captures the *strategic equilibrium reaction* of rational bidders to an important feature of FAs we refer to as the *FA curse* for its similarity with the winner's curse in common value auctions [Krishna, 2002]. The FA curse captures the fact that for the FA winner selling the good is in some sense 'bad news', as the selling event is positively correlated with the event in which costs are high. This is driven by the dependence of both the spot market price and the supplier's costs on $X$. In equilibrium, rational bidders respond to avoid the FA curse by charging the conditional expectation in (3.1), which is larger than the expected value of $X$. To illustrate the role of this correlation in the FA outcome we describe in Figure 3.3 the three different scenarios the FA winner may face depending on the outcome of $X$. For a given realization $c_j + z_j$, the spot market price is $X + c_j + z_j$. If the realization of $X$ is low enough, the spot market price is lower than the winning bid, and demand is satisfied in the spot market. For a moderate realization of $X$, the winning bid may be lower than the spot market price, but larger than the realized costs $c_i + X$; hence, the FA winner provides the product and makes positive profits. If the realization of $X$ is large enough, the FA winner provides the product with negative profits.

[7]This is a direct generalization of the result $E[X|X \geq a] \geq E[X]$ for some random variable $X$ and constant $a$.

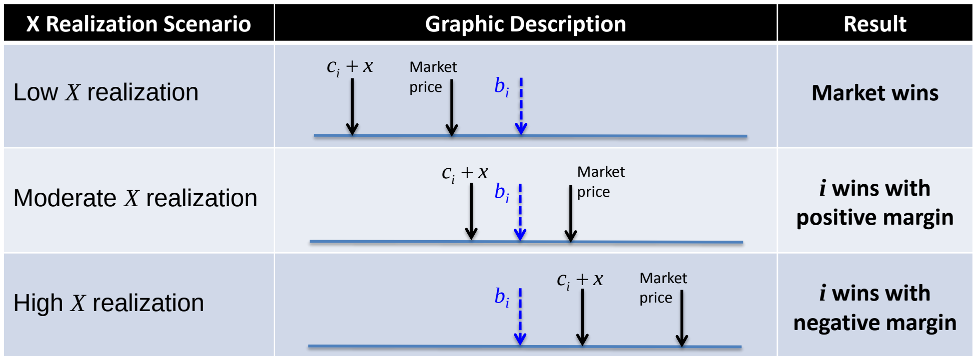



Figure 3.3: The “FA curse”.

The manner in which strategic suppliers react to the FA curse formalizes the observation by practitioners described in §1 that bid prices may be higher in a FA relative to a standard first-price auction run at the time the good is needed after observing the realization of $X$ (see Theorem 10 below). An important objective of the rest of the paper will be to derive mechanisms that alleviate the impact of the strategic response to the FA curse on buying prices.

**Concentrated Markets.** In concentrated markets we assume the number of suppliers in the open market is small, and thus bidders do not neglect the possibility of being sampled from the open market in the buying stage. Assuming a strictly increasing competitors' strategy $\beta$, the expected profit of supplier $i$ with private cost $c_i$ and bid price $b_i$ is given by:

$$\begin{aligned} \pi_i(b_i, c_i, \beta) \quad = \quad & \mathbb{E}\Bigg[\mathbb{I}\{b_i \leq \beta(c_{(1:N),-i})\}\left(\sum_{j=1, j\neq i}^{M} q_{j,M}\mathbb{I}\{b_i \leq c_j + X + Z_j\}\right)\cdot(b_i - c_i - X) \\ & + q_{i,M}\mathbb{I}\{b_i \leq \beta(c_{(1:N),-i})\}(\min\{b_i, c_i + X + Z_i\} - c_i - X) \\ & + q_{i,M}\mathbb{I}\{b_i > \beta(c_{(1:N),-i}) > c_i + X + Z_i\}Z_i\Bigg]. \end{aligned} \tag{3.2}$$

The first term captures the potential profit from winning the FA with a bid that is lower than the sampled spot market price. The second term captures profits captured by the FA potentially through the spot market; hence, revenues are the minimum between the bid submitted by $i$ at $t = 0$ and the spot market price of $i$ at $t = 1$. The third term captures profits from losing the FA, but being sampled at the spot market (with a price that lower than the winning bid).

**Proposition 8 (BNE bids in naive FA with concentrated markets)** *Suppose that $T = 1$. Let $\beta^{NFA}(\cdot)$ be a strictly increasing, continuous, and symmetric BNE strategy profile for the naive FA. Then, $\beta^{NFA}(\cdot)$ satisfies the following integral equation for all $c_i \in [\underline{c}, \bar{c}]$:*

$$\begin{aligned}\beta^{NFA}(c_i) &= c_i + \underbrace{\frac{\mathbb{E}\left[X\mathbb{I}\{c_i \le c_{(1:N),-i}\}\left(\sum_{j=1}^{M} q_{j,M}\mathbb{I}\{\beta^{NFA}(c_i) \le c_j + X + Z_j\}\right)\right]}{\mathbb{E}\left[\mathbb{I}\{c_i \le c_{(1:N),-i}\}\left(\sum_{j=1}^{M} q_{j,M}\mathbb{I}\{\beta^{NFA}(c_i) \le c_j + X + Z_j\}\right)\right]}}_{(A)} \\ &- \underbrace{\frac{\int_{c_i}^{\bar{c}} \frac{\partial \pi_i(\beta^{NFA}(c), c, \beta^{NFA})}{\partial c} dc}{\mathbb{E}\left[\mathbb{I}\{c_i \le c_{(1:N),-i}\}\left(\sum_{j=1}^{M} q_{j,M}\mathbb{I}\{\beta^{NFA}(c_i) \le c_j + X + Z_j\}\right)\right]}}_{(B)} \\ &+ \underbrace{q_{i,M}\frac{\mathbb{E}\left[Z_i\mathbb{I}\{\beta^{NFA}(c_{(1:N),-i}) > X + \bar{c} + Z_i\} - Z_i\mathbb{I}\{\beta^{NFA}(\min\{c_i, c_{(1:N),-i}\}) > c_i + X + Z_i\}\right]}{\mathbb{E}\left[\mathbb{I}\{c_i \le c_{(1:N),-i}\}\left(\sum_{j=1}^{M} q_{j,M}\mathbb{I}\{\beta^{NFA}(c_i) \le c_j + X + Z_j\}\right)\right]}}_{(C)}.\end{aligned}$$

Term $(A)$ captures the strategic reaction to the FA curse as before. Term $(B)$ is the information rent term considering that the FA winner may or may not be selected as the outside option in the spot market when the market is concentrated; the structure of the partial derivative is specified in the proof of the proposition in Appendix C.3. Term $(C)$ captures the *spot market opportunity effect*, representing additional margin the FA winner may get if selected from the spot market, while adjusting for the respective margin the least efficient supplier realizes. Because of the spot market opportunity effect, the analysis of the concentrated market regime is significantly more challenging than that of the diffused market. Hence, some of our results will specialize for the latter regime, which as we already argued, is by itself relevant in practice.

The procurement agency would like to minimize the *expected payment (or buying price)* corresponding to the overall sourcing cost, which considers the possibility to buy from the FA winner or from the spot market. In the naive FA the expected payment is:

$$\mathbb{E}\left[P^{NFA}\right] = \mathbb{E}\left[\min\left\{\beta^{\text{NFA}}(c_{i:N}), \sum_{j=1}^{M} q_{j,M}(c_j + X + Z_j)\right\}\right]. \tag{3.3}$$

### 3.3.2 Monitored FA

In the monitored FA, the price set in the auction stage is compared with the one charged by the FA winner in the open market and the buying price is set to be the lowest of the two, where demand is always satisfied by the FA winner. Assuming a strictly increasing competitors' strategy $\beta$, the expected profit of supplier $i$ with private cost $c_i$ and bid price $b_i$ is given by:

$$\pi_i(b_i, c_i, \beta) \;=\; \mathbb{E}\left[\mathbb{I}\{b_i \le \beta(c_{(1:N),-i})\} \cdot (\min\{b_i, c_i + X + Z_i\} - c_i - X)\right],$$

where $Z_i$ is the random markup that $i$ charges in the spot market. The expression above considers that bidder $i$ wins the FA if and only if he has the lowest bid, and that in that case, the transacted payment is the lowest between the bid submitted by $i$ at $t = 0$ and his spot market price at $t = 1$.

**Proposition 9 (BNE bids in monitored FA)** *Suppose that $T = 1$. Let $\beta^{MFA}(\cdot)$ be a strictly increasing, continuous, and symmetric BNE strategy profile for the monitored FA. Then, $\beta^{MFA}(\cdot)$ satisfies the following integral equation for all $c_i \in [\underline{c}, \bar{c}]$:*

$$\begin{aligned}\beta^{MFA}(c_i) \quad &= \quad c_i + \underbrace{\frac{\mathbb{E}\left[X\mathbb{I}\{\beta^{MFA}(c_i) \leq c_i + X + Z_i\}\right]}{\mathbb{E}\left[\mathbb{I}\left\{\beta^{MFA}(c_i) \leq c_i + X + Z_i\right\}\right]}}_{(A)} + \underbrace{\frac{\int_{c_i}^{\bar{c}} \bar{F}^{N-1}(c) \cdot \mathbb{E}\left[\mathbb{I}\{\beta^{MFA}(c) \leq c + X + Z_i\}\right] dc}{\bar{F}^{N-1}(c_i)\mathbb{E}\left[\mathbb{I}\{\beta^{MFA}(c_i) \leq c_i + X + Z_i\}\right]}}_{(B)} \\ &- \underbrace{\frac{\mathbb{E}\left[Z_i\mathbb{I}\{\beta^{MFA}(c_i) > \alpha(c_i + X + Z_i)\}\right]}{\mathbb{E}\left[\mathbb{I}\{\beta^{MFA}(c_i) \leq c_i + X + Z_i\}\right]}}_{(C)}.\end{aligned}$$

As previously, term $(A)$ captures the bidders' strategic reaction to the FA curse. Term $(B)$ is the information rent and term $(C)$ captures the market opportunity effect: the added value of delivering the good from the open market. In monitored FAs this value is captured by the FA winner and thus makes equilibrium bids more aggressive. The expected buying price is given by:

$$\mathbb{E}\left[P^{\mathrm{MF}A}\right] = \mathbb{E}\left[\min\left\{\beta^{\mathrm{MFA}}(c_{(i:N)}), c_{(i:N)} + X + Z_{(i:N)}\right\}\right], \tag{3.4}$$

where $Z_{(i:N)}$ is the markup charged in the open market by the lowest cost supplier in the FA. Comparing equations (3.3) and (3.4), it is immediate that the expected value of the right-hand side of the latter is smaller than the expected value of the right-hand side of the former; this is a direct effect of monitoring the market price of the most efficient FA participant. However, comparing the left-hand sides is not trivial as the equations for the BNE bidding strategies in Propositions 8 and 9 do not admit closed-form solutions. Further, in §6 we demonstrate through numerical experiments that equilibrium bids between monitored and naive FAs cannot be ordered in general. In §4 we propose a mechanism design approach that will more easily allow us to compare the expected buying prices between the different FAs and also with other benchmarks that we introduce below.

**Existence of Equilibrium.** Due to the correlation between the random common cost and the spot market price existence does not follow from standard first price auction existence results. In Appendix C.7 we study the existence of BNE bidding strategies in two cases of interest. Considering a finite action set, in Proposition C7 we prove the existence of symmetric BNE in increasing strategies for the naive FA in diffused markets and for the monitored FA, and this result is then extended to continuous and compact action sets. In Proposition C8 we show that in these settings any increasing symmetric BNE must be strictly increasing and identify sufficient conditions for it to be continuous. Similar results can be proved for the 'flexible FA' that will be introduced in §5. Our approach uses the techniques from Athey [2001].

## 3.4 Mechanism Design Approach

The bidding analysis in §3 was aimed at developing intuition regarding the various elements that affect bidding behavior of suppliers in FAs, with emphasis on the equilibrium response to the FA curse. In this section, we adopt a mechanism design approach to derive the expected buying price for various designs of FAs that were discussed in the previous sections, and that will more easily allow a comparison between them. In this section we come back to the original multi-period setting.

The analysis in this section has some novelties relative to existing mechanism design literature (e.g., Myerson 1981) as we now explain. Recall that the spot market price of supplier $i$ at time $t$ is given by $c_i + X_t + Z_{i,t}$. The price of the outside option is a function of these market prices; in the monitored FA it is the market price of the FA winner, and in the naive FA it is a random sample from all potential suppliers. Hence, the price of the outside option is a function of the *actual private cost realizations* $\boldsymbol{c}$ in addition to the realizations of $X_t$ and $\boldsymbol{Z}_t$. As a result, the allocation rules that determine whether a good is bought from the FA winner at the winning bid price depend on the private cost realizations of suppliers through the spot market prices. In other words, the outside option for the auctioneer is not exogenous but rather depends on the actual cost realizations, which differs from the classical mechanism design setup. This dependence yields additional terms when applying the envelope theorem to express the payments as a function of the allocation rule; despite this challenge, we derive useful expressions to compare expected buying prices.

### 3.4.1 Preliminaries

We consider the following class of first-price FA mechanisms that include the monitored and naive FAs as special cases. At time $t = 0$, the auctioneer receives bids from the FA participants and decides the FA winner. At time period $t = 1, 2, \ldots, T$, the auctioneer observes the realized spot market price and decides whether to buy from the FA winner or to buy from the spot market at period $t$. Formally, we define the allocation function $r_{i,t} : \mathcal{A}^N \times [\underline{p}_t, \bar{p}_t] \to [0,1]$, where $r_{i,t}(b_i, \boldsymbol{b}_{-i}, p_t)$ is the probability bidder $i$ sells the good through the FA if he submits a bid $b_i$, his competitors submit bids $\boldsymbol{b}_{-i}$, and the realization of the spot market price at period $t$ is given by $p_t$. The interval

$[\underline{p}_t, \bar{p}_t]$ is the range of feasible spot market prices. We denote $\boldsymbol{r}_t = (r_{1,t}, ..., r_{N,t})$ and $\boldsymbol{r} = (\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T)$ as the vectors of the allocation probabilities. In what follows we only consider allocation rules that satisfy $\sum_{i=1}^{N} r_{i,t}(\boldsymbol{b}, p) \leq 1$, for all $\boldsymbol{b}, p, t$.

Similarly, define the payment function $m_{i,t} : \mathcal{A}^N \times [\underline{p}_t, \bar{p}_t] \to \Re$, where $m_{i,t}(b_i, \boldsymbol{b}_{-i}, p_t)$ is the expected FA payment bidder $i$ receives for given bids $(b_i, \boldsymbol{b}_{-i})$ and a given realization of $p_t$. We let $\boldsymbol{m}_t = (m_{1,t}, ..., m_{N,t})$ and $\boldsymbol{m} = (\boldsymbol{m}_1, \ldots, \boldsymbol{m}_T)$. Because we consider first-price FA mechanisms, it may also be convenient to write $m_{i,t}(\boldsymbol{b}, p) = b_i r_{i,t}(\boldsymbol{b}, p)$, where $b_i$ is the bid submitted by bidder $i$ at $t = 0$. An FA mechanism is given by the functions $\boldsymbol{w} = (\boldsymbol{r}, \boldsymbol{m})$.

Let $A_j$ be the event in which bidder $j$ is selected in the spot market; when sampling randomly under the naive FA; $A_j = 1$ with probability $q_{j,M}$. Then, the spot market price is given by $p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t) = \sum_{i=1}^{M} A_i p_{i,t}(c_i, X_t, Z_{i,t})$, where $p_{i,t}(c_i, X_t, Z_{i,t})$ is the price charged at time $t$ by firm $i$ at the spot market. For the monitored FA, the allocation function is:

$$r_{i,t}(b_i, \boldsymbol{b}_{-i}, p_t) = \mathbb{I}\{b_i \leq b_j, \forall j \neq i\} \, \mathbb{I}\{b_i \leq p_{i,t}(c_i, X_t, Z_{i,t})\},$$

and for the naive FA, the allocation function is:

$$r_{i,t}(b_i, \boldsymbol{b}_{-i}, p_t) = \mathbb{I}\{b_i \leq b_j, \forall j \neq i\} \, \mathbb{I}\{b_i \leq p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)\}.$$

While to simplify the exposition we ignore ties, we note that all the results in this section hold even if ties are allowed. In both cases above the FA winner is determined by the lowest submitted bid; in the monitored FA the winning bid is compared against the FA winner's own spot market price, and in the naive FA it is compared against a randomly sampled spot market price. In both cases, the allocation function is equal to one if the good is allocated *through the FA* at the winning bid. However, FA winner's profits also consider the event of supplying the good through the spot market. For example, in the monitored FA, the allocation function is equal to one if the FA winner sells the good at his winning bid $b_i$. The FA winner profit function also considers the event of selling the good at his own spot market price when the latter is lower than $b_i$.

We benchmark the expected payments under the monitored FA and the naive FA against the expected payment in a procurement mechanism where a first price auction (FPA) is held at every time period $t$ after $X_t$ gets realized, and then the winner (the supplier with the lowest bid) immediately delivers the good. In these FPAs, bidders do not face cost uncertainty. In such a mechanism, the profit with bid $b_i$, cost $c_i$, and realization $x_t$ is given by:

$$\pi_{i,t}(b_i, c_i, \beta) = \mathbb{P}\{b_i \leq \beta(c_j), \forall j \neq i\} \cdot (b_i - c_i - x_t).$$

We denote the payment in the mechanism in which a FPA is run at every time period by $P^{FPA}$. In a BNE, the expected payment for the FPA at period $t$ is given by:

$$\mathbb{E}\left[P_t^{FPA}\right] = \mathbb{E}\left[\beta_{N,t}^{FPA}(c_{(1:N)})\right] = \mathbb{E}\left[c_{(1:N)} + X_t + F(c_{(1:N)})/f(c_{(1:N)})\right], \tag{3.5}$$

where $c_{(1:N)}$ is the lowest order statistic among the private costs. The function $\beta_{N,t}^{FPA}(\cdot)$ is the BNE strategy of the FPA with $N$ bidders. The second equality holds by standard mechanism design arguments based on the envelope theorem [Milgrom, 2004]. In the next subsection, we compare the payments in the naive FA with those of FPA to quantify the impact of the equilibrium response to the *FA curse*.

A natural alternative benchmark one may consider is a first price auction with a random reserve price that is set to be equal to the spot market price (randomly sampled like in the diffused market naive FA); in Appendix C.4 we show that expected payments under the two FPAs (with and without the reserve price) are asymptotically equivalent as the number of bidders grows large. Hence, the comparison result shown in the following subsection also holds for FPA with an outside option that is given by the spot market price.

### 3.4.2 The Naive FA

We next provide a lower bound for expected payments in the naive FA that only depends on the allocation rule and will allow for easier comparisons with other mechanisms.

**Proposition 10 (Bound on the expected buying price in naive FA)** *Let $\boldsymbol{\beta}(\cdot)$ be a strictly increasing symmetric BNE strategy profile induced by a naive FA mechanism $\boldsymbol{w} = (\boldsymbol{r}, \boldsymbol{m})$. Then, the expected buying price for the auctioneer can be lower bounded by the sum over t of the following expression, $\mathbb{E}\left[P^{NFA}\right] \geq \sum_{t=1}^{T} P_t^{NFA}$, with*

$$P_t^{NFA} = \sum_{i=1}^{N} \pi_{i,t}(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}) + \mathbb{E}\left[\widetilde{q}_0(\boldsymbol{c}, X_t)\right] + \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t)\left(v(c_i) + X_t - \widetilde{q}_0(\boldsymbol{c}, X_t)\right)\right],$$

*where $\widetilde{q}_0(\boldsymbol{c}, X_t) = \sum_{i=1}^{M} A_i(c_i + X_t)$, and $v(c_i) = c_i + \frac{F(c_i)}{f(c_i)}$ is the virtual cost.*

We next compare the expected payments under naive FA with those obtained under FPA. To derive the result we consider an asymptotic regime in which the number of participant firms in the FA is large. This simplifies the analysis as it allows to get a better handle on the FA equilibrium bids (which are complex objects without closed-form expressions for every finite number of participants), yet leads to meaningful results as expected payments can be strictly ordered.

**Theorem 10 (Asymptotic performance of naive FA)** *Assume* $\min_{c\in[\underline{c},\bar{c}]} f(c) > 0$*. Then:*

1. (Comparison to first price actions in concentrated markets)

$$\lim_{N\to\infty} \mathbb{E}\left[P^{NFA} - P^{FPA}\right] \geq 0.$$

2. (Comparison to first price actions in diffused markets) *If* $\underline{z} + \underline{x} < \mathbb{E}[X_t]$ *for all* $t$*, and the spot market is diffused, then:*

$$\lim_{N\to\infty} \mathbb{E}\left[P^{NDFA} - P^{FPA}\right] > 0.$$

In §6 we provide numerical results for small values of $N$ that support the validity of the asymptotic results for small markets. As running a first price auction after the common cost is realized eliminates the FA curse associated with cost uncertainty, Theorem 10 demonstrates the impact of bidders' strategic response to the FA curse on expected payments. In concentrated markets, the expected payments of the naive FA are larger than those of the FPA, where this inequality is strict in diffused markets. Note that governments may still prefer to use FAs because running auctions as needs arise has administrative costs associated to setting up the auctions, receiving bids, processing them, and so forth, that scale with the total number of demand periods. In contrast, when running an FA the administrative cost associated to running it is only payed once, at $t = 0$.

### 3.4.3 The Monitored FA

We next demonstrate that one may improve the performance of FAs by monitoring the spot market price of the FA winner. We first derive an expression for the expected payments in monitored FAs.

**Proposition 11 (Expected buying price in monitored FA)** *Let* $\boldsymbol{\beta}(\cdot)$ *be a strictly increasing symmetric BNE strategy profile induced by a monitored FA mechanism* $\boldsymbol{w} = (\boldsymbol{r}, \boldsymbol{m})$*. Then, the expected total buying price for the auctioneer from period 1 to period* $T$ *is given by:* $\mathbb{E}[P^{MFA}] = \sum_{t=1}^{T} P_t^{MFA}$*, where*

$$P_t^{MFA} = \mathbb{E}\left[q_0(\boldsymbol{c}, X_t) + \sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t))\,(v(c_i) + X_t - q_0(\boldsymbol{c}, X_t))\right],$$

*with* $q_0(\boldsymbol{c}, x) = c_{(1:N)} + x$*,* $v(c_i) = c_i + \frac{F(c_i)}{f(c_i)}$*, and* $c_{(1:N),-i}$ *is the lowest order statistic among the cost realizations of firm i's competitors.*

In Proposition 11 we derive an expression for the expected payment that depends only on the allocation rule, even though the allocation rule itself also depends on the actual realizations of costs through the spot market price. A key idea in this derivation is to exploit the fact that some of the additional terms appearing in the envelope expression cancel out. In what follows we leverage this result to show that when the number of suppliers is large the monitored FA is payment equivalent to running FPAs as needs arise. As a corollary from this result and Theorem 10 we obtain an order between the expected payments under monitored FA and naive FA.

**Theorem 11 (Asymptotic performance of monitored FA)** *Assume* $\min_{c\in[\underline{c},\bar{c}]} f(c) > 0$.

1. (Payment equivalence to first price auctions)
$$\lim_{N\to\infty} \mathbb{E}\left[P^{MFA} - P^{FPA}\right] = 0.$$
2. (Comparison to naive FA in concentrated markets)
$$\lim_{N\to\infty} \mathbb{E}\left[P^{NFA} - P^{MFA}\right] \geq 0.$$
3. (Comparison to naive FA in diffused markets) *Assume* $\underline{z} + \underline{x} < \mathbb{E}[X_t]$ *for all* $t$*, and that the spot market is diffused under naive FA. Then:*
$$\lim_{N\to\infty} \mathbb{E}\left[P^{NDFA} - P^{MFA}\right] > 0.$$

Theorem 11 shows that the monitored FA alleviates the impact of bidders' reaction to the FA curse. This conclusion might be somewhat intuitive: the monitored FA should provide lower expected buying prices relative to the naive FA, because it uses a better outside option (the FA winner spot market price as oppose to a randomly sampled spot market price). However, the analysis is not direct, because the expected buying price also depends on the equilibrium bids, which are hard to characterize. By taking $N \to \infty$ we can simplify the dependence of the expected buying prices on equilibrium bids and prove the result. In §6 we provide numerical analysis demonstrating that the monitored FA typically performs better than the naive FA even when $N$ is small.

A key for establishing Theorems 10 and 11 is to show that the per period asymptotic expected payment under the monitored FA and first-price auction are no larger than $\underline{c} + \mathbb{E}[X_t]$. This follows because as $N$ grows the lowest cost supplier achieves $\underline{c}$ and competition dissipates all markups. In contrast, one can also show that the expected payment under the naive FA is strictly larger than this quantity due to the strategic response to the FA curse, even in the regime of large $N$.

## 3.5 Price Flexibility in FAs

The results we have established so far suggest that monitoring spot prices of FA winners may improve the performance of FAs. Nevertheless, as discussed before such an approach involves practical challenges (such as centralized monitoring costs) and therefore may not be necessarily followed by procurement agencies (as is the case for ChileCompra). In this section we discuss practical alternative approaches to improve the performance of FAs. We study FA designs that alleviate the common cost uncertainty and therefore achieve lower payments relative to naive FAs. We show that the optimal mechanism eliminates the common cost uncertainty by indexing payments to the common cost. One limitation of using this type of mechanisms in practice, though, is that it may be hard to establish prices indexes for many of the goods and services procured through FAs (e.g., computers, office equipment, services). Thus motivated, we introduce practical variants of the naive FA design, a class of *flexible FAs*, which under some conditions achieve payments that are similar to the ones under the optimal mechanism and the monitored FA, without requiring a price index nor centralized monitoring of spot market prices.

To facilitate the analysis we focus on diffused markets, where the price of the outside option at each period $t$ is given by $p_t = \sum_{j=1}^{M} A_j(c_j + X_t + Z_{j,t})$. It is simple to show that $p_t$ converges in distribution to $X_t + Z_t$ as $M \to \infty$, where we abuse notation assuming that $Z_t$ has the same distribution as $c_j + Z_{j,t}$. (Recall that in diffused markets $Z_t$ is independent of $c_i$, for all $i \in \mathcal{N}$.)

### 3.5.1 The Flexible and Restricted-Flexible FA

One proxy for the realization of the common cost (which in general, is not directly observed by the procurement agency) is the realization of the spot market price. Hence, a practical FA variant may use observed spot market prices to partially remove common cost uncertainty. We introduce a class of *flexible FAs*, in which at every point in time the FA winner *has the option* to *decrease* his bid to match the observed spot market price (in case the latter is lower than the winning bid). When lowering price to match the spot market price the FA winner is guaranteed to supply the good at that period. We study two practically relevant variants of this new class of FAs.

First, we consider a *flexible FA* in which the decision to lower the price is done independently each period. In this case, the FA winner can match the spot market price one period, but then the bid price goes back up to its original value at the next period. A flexible FA may result in a wiggly time-series of payments to the FA winner. Discussions with ChileCompra suggested that for transparency reasons it could be preferred that the FA winner does not increase prices during the FA time horizon. In fact, in practice it is often the case that if the FA winner reduces its price for one buyer, then the procurement agency expects the supplier to make the reduced price available for all future buyers as a way of encouraging transparency and competitive prices. For these reasons, an appealing variant is a restricted-flexible FA, in which once the FA winner reduces his price to match the spot market price, he cannot increase it back again. In this case, payments to the FA winner are always non-increasing, which is an appealing practical feature.

Note that adopting a multi-period model is key to study the restricted-flexible FA, because it induces the following dynamic trade-off for the FA winner that needs to be taken into account. On one hand, matching the spot market price today may increase current profits. On the other hand, since the FA winner would not be able to increase the price back in future periods, matching may reduce future profits (and even cause losses) when realized costs are higher. Note that under a flexible FA optimal matching decisions are myopic and such a tradeoff does not exist.

Now, we formalize expressions for expected payments. The allocation rule of the flexible FA for time period $t$ is given by:

$$r_{i,t}(b_i, \boldsymbol{b}_{-i}, z_t, x_t, c_i) \; = \; \mathbb{I}\{b_i \leq b_j, \; j \neq i\} \cdot \left( \mathbb{I}\left\{b_i \leq z_t + x_t\right\} + \mathbb{I}\left\{b_i > z_t + x_t, c_i \leq z_t\right\}\right).$$

In this allocation rule the FA winner sells through the FA either when his bid is lower than the spot market price or when he can afford matching the spot market price (the latter is the case if $c_i \leq z_t$). The payment function of the flexible FA for time period $t$ is given by:

$$m_{i,t}(b_i, \boldsymbol{b}_{-i}, z_t, x_t, c_i) \; = \; \mathbb{I}\{b_i \leq b_j, \; j \neq i\} \cdot \left( b_i \mathbb{I}\left\{b_i \leq z_t + x_t\right\} + (z_t + x_t)\mathbb{I}\left\{b_i > z_t + x_t, c_i \leq z_t\right\}\right).$$

Even though the allocation depends on the *actual cost realizations*, using arguments similar to the ones used for the monitored FA, we establish the following expected payment characterization.

**Proposition 12 (Expected buying price in Flexible FA)** *Let $\boldsymbol{\beta}(\cdot)$ be a strictly increasing symmetric BNE strategy profile induced by the flexible FA. Then, the expected payment in flexible FAs is given by: $\mathbb{E}[P^{FLE}] = \sum_{t=1}^{T} P_t^{FLE}$, where*

$$P_t^{FLE} = \mathbb{E}[X_t + Z_t] + \mathbb{E}\left[\left(v(c_{(1:N)}) - Z_t\right) \cdot \left(\mathbb{I}\left\{c_{(1:N)} \leq Z_t\right\} + \mathbb{I}\left\{\beta(c_{(1:N)}) \leq Z_t + X_t, c_{(1:N)} > Z_t\right\}\right)\right].$$

We now turn to consider the restricted-flexible FA.[8] Suppose that at time period $t$, the FA winner current bid price is $b_t$, the private cost is $c$, and $x_t$ and $z_t$ are realized. Then, the FA winner faces a dynamic optimization problem with the following Bellman equation:

$$\tilde{V}_t(b_t, c, x_t, z_t) = \begin{cases} \max\{V_{t+1}(b_t, c, x_t),\ (z_t - c) + V_{t+1}(x_t + z_t, c, x_t)\}, & \text{if } b_t > x_t + z_t \\ b_t - c - x_t + V_{t+1}(b_t, c, x_t), & \text{otherwise,} \end{cases} \tag{3.6}$$

where $V_t(\cdot)$ is the value to go and $V_{t+1}(b_{t+1}, c, x_t) = \mathbb{E}_{x_t}\left[\widetilde{V}_{t+1}(b_{t+1}, c, X_{t+1}, Z_{t+1})\right]$. When $b_t > x_t + z_t$ the FA winner has the option to match the spot market price at time period $t$, and does so if and only if $V_{t+1}(b_t, c, x_t) \leq (z_t - c) + V_{t+1}(x_t + z_t, c, x_t)$. In Appendix C.5 (see Proposition C3), we show that $V_{t+1}(b_t, c, x_t)$ is strictly increasing in $b_t$; hence, the FA winner matches if current realized profits $(z_t - c)$ compensate for the decrease on future profits as $b_t$ decreases to $x_t + z_t$. Using the dynamic programming formulation and a mechanism design approach, we also provide in Appendix C.5 an explicit expression for the expected total payments under the restricted-flexible FA along the entire FA horizon, which we denote by $P^{FLR}$.

In the following subsection we leverage these expressions to compare the expected payments of these flexible FAs with those of previously studied mechanisms.

### 3.5.2 Expected Payments for Flexible FAs

To evaluate the flexible FA we benchmark it against three mechanisms: the monitored FA; the FPA, in which an auction is run at every time period without cost uncertainty; and a mechanism that is optimal in a class of mechanisms that generalizes the class defined in §3.4. While the latter is described in detail in Appendix C.6, in what follows we summarize the setup and main results.

We define the allocation function $r_{i,t} : \mathcal{A}^N \times [\underline{x}, \bar{x}] \times [\underline{z}, \bar{z}] \to [0, 1]$, where $r_{i,t}(b_i, \boldsymbol{b}_{-i}, x, z)$ is the probability bidder $i$ sells the product in period $t$ if he submit a bid $b_i$, his competitors submit bids $\boldsymbol{b}_{-i}$, the common cost realization is $x$, and the realization of $Z$ is $z$. We let $\boldsymbol{r}_t = (r_{1,t}, ..., r_{N,t})$ and $\boldsymbol{r} = (\boldsymbol{r}_1, ..., \boldsymbol{r}_T)$. We consider allocation rules that satisfy $\sum_{i=1}^{N} r_{i,t}(\boldsymbol{b}, x, z) \leq 1$, for all $\boldsymbol{b}, x, z$,and $t$. Similarly, define the payment function $m_{i,t} : \mathcal{A}^N \times [\underline{x}, \bar{x}] \times [\underline{z}, \bar{z}] \to \Re$, where $m_{i,t}(b_i, \boldsymbol{b}_{-i}, x, z)$ is the expected payment bidder $i$ receives in period $t$ for given bids $(b_i, \boldsymbol{b}_{-i})$ and given realizations $x$ and

[8]For technical simplicity, for the analysis of the restricted-flexible FA we assume that the process $\{X_t;\ t \geq 1\}$ is Markov and that $\{Z_t;\ t \geq 1\}$ are i.i.d.. We could generalize this assumption to the following: the distribution of $(Z_t, X_t)$ depends only on the realization of $X_{t-1}$, and $X_t$ and $Z_t$ are independent conditional on $X_{t-1}$.

$z$. We let $\boldsymbol{m}_t = (m_{1,t}, ..., m_{N,t})$ and $\boldsymbol{m} = (\boldsymbol{m}_1, ..., \boldsymbol{m}_T)$. An FA mechanism is given by the functions $\boldsymbol{w} = (\boldsymbol{r}, \boldsymbol{m})$. This class is more general than the one defined in §3.4, as it allows dependence on $x$ and $z$ separately, beyond just $p = x + z$ (but it does not include the flexible FA because the allocation rule of the latter depends on the cost realization).

In Appendix C.6 (see Proposition C6) we characterize the optimal mechanism as a "modified" second price auction, in which the winner is payed $b_{(2)} + x_t$ (where $b_{(2)}$ denotes the second lowest bid), with an appropriately chosen random reserve price (that depends on $z_t$). Note that under the optimal mechanism derived in Appendix C.6, payments in every period $t$ are indexed to the realized common cost $x_t$, effectively creating a price index that eliminates the common cost uncertainty for the FA winner, and therefore, eliminating the FA curse.

It is important to observe that while the practicality of the price index FA and the monitored FA lies on the availability of relevant price indexes and centralized monitoring operations, the flexible FA does not require any of these. In fact, this new class is based on the same decentralized sampling of spot market prices that is used in the naive FA, and only requires giving the FA winner the additional option of matching the spot market price at every time period. The following result shows that the expected buying price under a flexible FA is asymptotically optimal, where we denote $P^{OPT}$ as the expected payments over the horizon under the optimal mechanism. Furthermore, the expected payments under the flexible FA asymptotically coincide with those of the monitored FA, as well as running first price auctions as needs arise.

**Theorem 12 (Asymptotic optimality of the flexible FA)** *Assume that* $\min_{c\in[\underline{c},\bar{c}]} f(c) > 0$. *Then,*

$$\begin{aligned} \lim_{N\to\infty} \mathbb{E}\left[P^{FLE} - P^{OPT}\right] &= 0, \\ \lim_{N\to\infty} \mathbb{E}\left[P^{FLE} - P^{MFA}\right] &= 0, \\ \lim_{N\to\infty} \mathbb{E}\left[P^{FLE} - P^{FPA}\right] &= 0. \end{aligned}$$

In contrast with Theorems 10 and 11 that show that payments under the naive FA exceed those of the FPA and the monitored FA for large $N$, Theorem 12 implies that the flexible FA achieves low expected payments when the number of bidders is large, comparable to those obtained by the same benchmarks. Moreover, in the numerical analysis that is be described in §3.6 we show that expected payments under the flexible FA are typically close to these benchmarks even for small values of $N$.

The next theorem shows that in contrast, the restricted-flexible FA achieves expected payments that are larger or equal than those of the flexible FA when $N$ is large.

**Theorem 13** *Assume that* $\min_{c\in[\underline{c},\bar{c}]} f(c) > 0$*. Then:*

$$\lim_{N\to\infty} \mathbb{E}\left[P^{FLR} - P^{FLE}\right] \geq 0.$$

Theorem 13 implies that even though the restricted-flexible FA is appealing in the sense that it induces non-increasing payments to the FA winner over the FA time horizon, this restriction may drive higher overall expected payments to the procuring agency when the number of suppliers is large. The intuition is as follows. Typically, the flexible FA induces lower equilibrium bids, while the restricted-flexible FA guarantees that once the FA winner matches the spot market price, the buying price is reduced from there on. When $N$ is large, however, bids are low to start with, and therefore, the FA winner is less likely to match the spot market price in the restricted-flexible FA; as a consequence, the flexible FA results in lower expected payments. In the following section we conduct numerical experiments exploring the implications of the restricted-flexible FA for small values of $N$. In this case, we show that the restricted-flexible FA achieves comparable, or sometimes even better, expected payments relative to the flexible FA, because the FA winner matches the spot market more often.

## 3.6 Numerical Experiments

We complement the analytical results obtained in the previous sections with numerical experiments demonstrating the robustness of the former. Two of the main insights we have derived so far for the design of FAs are that (1) the flexible FA achieves lower expected payments relative to the naive FA, and that (2) the performances of the flexible FA and the monitored FA are comparable. While the main analytical results deriving these insights (in particular Theorems 11 and 12) are of asymptotic nature, our numerical results validate these conclusions even when the number of suppliers is relatively small. Our results also demonstrate that the expected payments under the restricted-flexible FA may be similar and even lower than the ones achieved under the flexible FA, despite the asymptotic result in Theorem 13.

**Setting.** To simplify the setting we assume throughout this section that $Z_{i,t} = \Delta$ is a deterministic parameter. We assume a diffused market, and set the spot market price to be $c_0 + X_t + \Delta$ where

$c_0$ is an exogenous parameter capturing different sampling mechanisms as we explain later. In the monitored FA the spot market price is set according to the private cost of the FA winner. Further, we assume that the common costs $\{X_t : t \geq 1\}$ are i.i.d. random variables.

**Price indexes and flexible FA.** Suppose that at some $t > 0$ the procuring agency can observe the realized common cost $x_t$. When this is the case (for example, when the common cost component is a commodity, such as gas), the buyer might use a perfect price index to eliminate the common cost uncertainty. Using such a price index, when the realized common cost is $x_t$, the payment of a winning bidder with bid $b^*$ can be set to $b^* + (x_t - \mathbb{E}[X_t])$, in which case costs for the FA winner are predetermined a priori and equal to $c + \mathbb{E}[X_t]$. One can also show that such "perfect price index FA" with an appropriate reserve price is optimal among the class of mechanisms introduced in §5 (that is, in the setting of this section, the optimal mechanism can be implemented as a first-price perfect price index FA). Further, under the assumptions of the present section, the flexible FA achieves the same expected payments as the perfect price index FA, and as running a first price auction after $X_t$ has been realized (FPA). Hence, the results obtained below for the flexible FA *also apply to FPA* as well as to the first-price perfect price index FA. We next describe the methodology and setup of our numerical experiments, followed by the main results and key insights.

### 3.6.1 Methodology

We numerically study the different FA models: naive FA, monitored FA, restricted-flexible FA, and flexible FA (which, as we mentioned, is payoff equivalent to FPA). One may observe that under the assumptions above, in the naive FA, monitored FA, and flexible FA it is sufficient to analyze the expected buying prices at a single period. In the restricted-flexible FA, however, outcomes are not independent across periods. To simplify the numerical solution of the dynamic program associated with the FA winner's decision problem we assume $T = 2$. Methodological background for the experiments is given in Appendix C.2. In particular, Proposition C1 establishes the ordinary differential equations (ODEs) and the corresponding boundary conditions that characterize symmetric BNE strategies for the various FAs. Similarly to asymmetric first-price auctions, our ODEs are not well-behaved at the boundary. Hubbard and Paarsch [2011] and Fibich and Gavish [2011] provide useful summaries of the challenges involved in numerically solving ODEs to establish BNE strategies in asymmetric first-price auctions, as well as ways to overcome these technical challenges (see also further discussion in Appendix C.2). We solve the ODEs using a Runge-Kutta method in Matlab to obtain the BNE bid functions for the various FA mechanisms.

For various instances described below and the different FA mechanisms, we compute the equilibrium bid functions using the ODEs, and then we use simulation to determine the expected buying prices. To do so, we randomly generate the private costs $c_i$ for all suppliers and use the BNE bid functions to obtain the winning bid $\beta(c_{(1)})$, where $c_{(1)}$ is the lowest realized private cost. Then, we randomly generate a common cost $X_t$. The realized cost of the winning bidder is $c_{(1)} + X_t$, and the realized spot market price is $c_0 + \Delta + X_t$ for naive FA with diffused market (the parameter $c_0$ can be associated with the level of competitiveness in the open market, or with different sampling mechanisms, as we further discuss below), flexible FA and restricted-flexible FA, and $c_{(1)} + \Delta + X_t$ for monitored FA. Given these quantities, buying prices are determined according to the rules of the different mechanisms. To simulate the expected buying price, we replicate the above procedure $50,000$ times for each problem instance and each FA mechanism; relative errors in the results below are all less than 0.5% with 98% confidence intervals.

### 3.6.2 Scenarios and Results

We consider $N$ bidders, where bidder $i$ has a private cost $c_i \sim U\left[0, \frac{1}{2}\right]$. We assume a uniform distribution in order to simplify the ODEs; we fix this distribution while varying other parameters. We consider a common cost $X_t$ that is distributed according to a truncated normal distribution with mean $\mu_x$, standard deviation $\sigma_x$, and support $[0, 2\mu_x]$; we denote by $\sigma_{max} = \frac{2\mu_x}{\sqrt{12}}$, the value of the standard deviation under which the distribution of $X_t$ becomes uniform over $[0, 2\mu_x]$. We consider all possible combinations of a large range of parameter values that capture different settings: $N \in \{2, 4, 6, 8, 10\}$, $c_0 \in \left\{\frac{1}{8}, \frac{1}{6}, \frac{1}{4}\right\}$, $\Delta \in \{\frac{1}{16}, \frac{1}{12}, \frac{1}{8}\}$, $\mu_x \in \left\{1, \frac{3}{2}, 2\right\} \times (c_0 + \Delta)$, and $\sigma_x \in \left\{\frac{\sigma_{max}}{2}, \frac{3\sigma_{max}}{4}\right\}$. In total, there are 270 model instances. Note that the different values of $c_0$ can represent different sampling mechanisms from the spot market: $c_0 = 1/4$ corresponds to the expected cost when sampling one supplier from the spot market, $c_0 = 1/6$ to the expected cost when sampling two and choosing the minimum, and $c_0 = 1/8$ when sampling three and choosing the minimum.

All numerical results can be obtained from the authors upon request. The plot in the left side of Figure 3.4 shows the BNE bid functions under the different mechanisms for one representative instance. The bid functions in the plot are ordered in the "typical order", which repeated itself

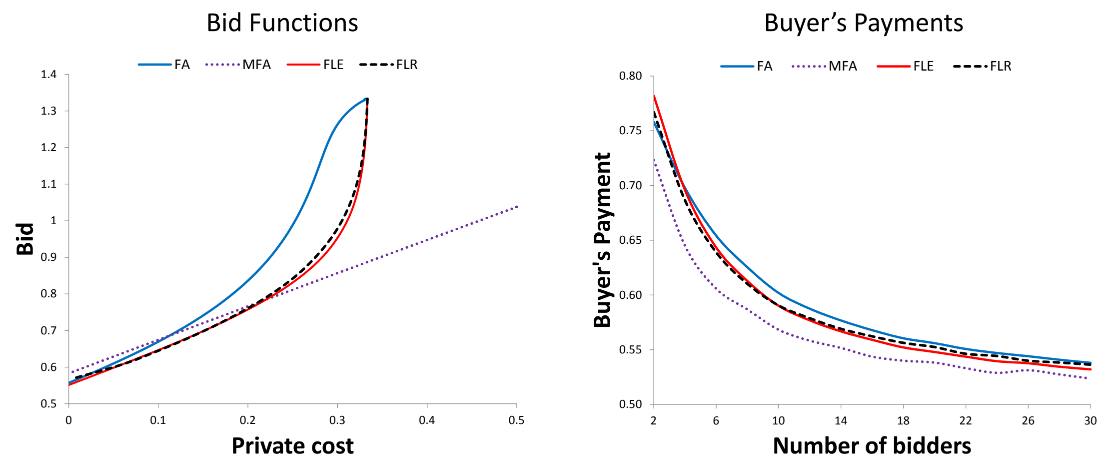



Figure 3.4: **A representative problem instance.** Parameter values are: $\mu_x = 0.5$, $\sigma_x = \frac{\mu_x}{\sqrt{12}} = \frac{\sigma_{max}}{2}$, $c_0 = 0.25$, and $\Delta = \frac{1}{12}$. (*Left*) BNE bid function under the NDFA, MFA, FLE, and FLR mechanisms with $N = 10$ bidding suppliers. (*Right*) Per-period expected buyer's payment for different mechanisms as a function of the number of participating suppliers. Because of payoff equivalence, the flexible FA curve also represents FPA and the perfect price index FA.

over most instances. The equilibrium bid under naive FA is higher than the equilibrium bid under flexible and restricted-flexible FAs. There is no universal order between equilibrium bids under naive FA and monitored FA.[9] The plot in the right hand side of Figure 3.4 depicts the buyer's expected payment per-period as a function of the number of bidders, illustrating that the established comparisons in Theorems 10, 11, and 12 may also hold under a relatively small number of bidders. In the representative example that is showed in the figure, monitored FA dominates naive FA for any number of bidders, while flexible and restricted-flexible FAs dominate naive FA for any $N \geq 5$, and approach the monitored FA as the number of bidders increases.

We find that in a diffused market with $c_0 = 1/4$, the buyer's expected payment under monitored FA is always smaller than the payment under naive FA for all scenarios, and that on average the prices are 9% lower under monitored FA. Note that this value of $c_0$ represents a sampling of an average cost seller from the spot market. However, when the spot market is cheaper than the average seller (that is, $c_0 < 1/4$), there is a small number of instances for which the expected payment under monitored FA is larger than that of naive FA. However, it is still the case that in most instances the

[9]We note that the FA winner under the naive FA, flexible FA, or restricted-flexible FA is competing against an outside market with price $c_0 + \Delta + X$, and as a result bidders with private cost higher than that value never win.

monitored FA performs better, especially when $X_t$ is volatile (with a larger coefficient of variation), and when there is more competition (when the number of sellers is relatively large). For example, with $c_0 = 1/6$ and $\sigma_x = 3\sigma_{max}/4$, the expected payments of the monitored FA are lower than those of the naive FA in 31 out of 36 parametric combinations of $\Delta$, $\mu_x$, and $N \geq 4$. These findings represent the following insight that is consistent with Theorem 11: monitored FAs typically perform better than naive FAs, except in some cases in which the number of FA participants is small and relatively efficient suppliers are sampled from the spot market in the naive FA.

Comparing the performance of flexible FA with that of naive FA, we observe that flexible FAs often dominate naive FAs under sufficient competition. For example, in 88% of the tested instances (191 of 225) in which $N \geq 4$, flexible FA outperformed naive FAs with a buying price that was on average 3.7% lower than the one obtained under the naive FA. In cases where in addition in a naive FA an "average cost" supplier was sampled ($c_0 = 1/4$), the flexible FA dominated the naive FA in an even more consistent manner (in more than 95% of these instances the payments under the flexible FA were lower). Note that a naive FA may perform better than a flexible FA when $N$ is small, because in naive FAs bidders face the additional competition of the random spot market price (which in flexible FAs can always be matched), incentivizing suppliers to bid more aggressively. This effects get dissipated quickly as $N$ grows.

Comparing the flexible and the monitored FAs, we observe that the competition level that is required for payments under a flexible FA to approach those under a monitored FA depend on the cost efficiency of the outside option sampling, which is captured by the value of $c_0$. When $c_0$ has a low value, we found that the performance of flexible FAs is close to that of monitored FAs even under moderate competition (for example, considering the cases of $c_0 \in \{1/6, 1/8\}$, when $N \geq 8$, monitored and flexible FAs have similar performances, except in 3 out of 72 instances). However, when average suppliers are sampled ($c_0 = 1/4$), monitored FAs outperformed flexible FAs in most of the instances that were tested even for $N = 10$. In such cases flexible FAs seem to require more aggressive competition (larger $N$) in order to achieve a performance comparable to monitored FAs. (See the right hand side of Figure 3.4 for one such instance.)

Comparing the flexible and the restricted-flexible FAs, we observe that the two mechanisms perform very similarly in terms of both bid functions as well as expected payments in our scenarios with $N \leq 10$. In contrast with the asymptotic result in Theorem 13, the average expected payments over our instances are slightly lower under the restricted-flexible FA relative to the flexible FA. Generally, the restricted-flexible FA outperforms the flexible FA in scenarios in which the FA winner is likely to reduce the winning bid under the former, resulting in a lower second period bid price. For small values of $N$ margins are relatively high and therefore the FA winner is likely to match the spot market in the first period under the restricted-flexible; this is specially true for low values of $c_0$. Because all instances are with $N \leq 10$, this is also consistent with the plot depicted on the right side of Figure 3.4, suggesting that the superiority of the flexible FA over the restricted-flexible FA in this setting becomes more apparent only for large $N$, and that for small $N$ they are similar and restricted-flexible FA may even outperform the flexible FA.

## 3.7 Conclusions

**Summary.** In this paper we introduced a novel auction model to analyze FAs, procurement mechanisms that are commonly used around the world. The increasing popularity of FAs can be attributed, among other things, to the following features: (i) they avoid the administrative costs of running repeated first-price auctions; and (ii) they tend to screen more cost-efficient suppliers relative to a setting in which each public organization buys from the spot market. Despite these advantages, we show that FAs are subject to a sort of winner's curse that makes expected buying prices higher relative to running first-price auctions when needs arise.

**Design Recommendations.** Based on our analysis we suggest the following prescriptions for a practical design of FAs that alleviate cost uncertainties and reduce buying prices:

1. Monitoring the price that is charged by the FA winner in the open market and using it to upper bound the buying price significantly reduces procurement costs. Procurement agencies should engage in this type of monitoring activities in cases where the associated administrative costs are not too high.

2. Building and implementing perfect price indexes for the common random parts of costs may reduce buying prices. The potential value of implementing price indexes suggests that procurement agencies should try to be creative in building such indexes even when there are not obvious ways of doing so, for example, when buying services or non-commodities products.

3. Allowing FA winners to reduce prices to match spot market prices through flexible FAs may be an effective way of reducing buying prices even when suppliers are restricted not to increase the price back after matching the spot market price. Flexible FA is especially effective when several (more than 2-3) suppliers participate in the FA.

While some of these prescriptions are intuitive, they need to be taken with some caution. For example, when the number of FA participants is small and the outside option given by the spot market is attractive, the monitored and flexible FAs may perform worse than naive FAs. This concern not withstanding, the above prescriptions are likely to be relevant in many cases of practical interest and have been considered by the Chilean government to improve their procurement processes. We hope that in the future they will also be considered by other buying agencies.

**Future work and open questions.** Our work is the first that provides a formal understanding on how the cost uncertainty suppliers face in FAs affects their bidding behavior and the outcomes in FAs. We focus on this aspect of FAs but several interesting directions may be worth exploring in the future. In particular, throughout the paper we assumed risk neutral bidders, and even in this case, the discussion regarding cost uncertainty was rich and relevant. An interesting extension of our model would be to consider risk averse bidders. This would introduce additional complications in the analysis; for example, under risk aversion revenue equivalence-style arguments may no longer hold. On the other hand, introducing risk averse bidders would allow a meaningful discussion on bidders' use of financial instruments such as options to hedge the risk associated to the common cost. We conjecture that the use of options may play a similar role to a perfect price index in the case of risk neutral bidders to alleviate the extent of the FA curse.

We have abstracted away from some problem aspects that would become relevant under demand uncertainty. One natural feature is non-linear suppliers' costs, capturing their economies of scale. A related interesting direction is the possibility of pre-commitment decisions suppliers may need to take before demand is realized. These considerations open a whole new set of issues worth considering in future work. Overall, we hope that this paper, together with the follow up work it may generate, will improve the way FAs and other procurement processes are designed in practice.

# Chapter 4

# On Preferences for Contractual Forms in Supply Chains

Lijian Lu and Yaozhong Wu

## 4.1 Introduction

Contracts are commonly used in supply chains to coordinate the activities of parties whose local objectives are not always perfectly aligned with one another. It is now well known that the wholesale price–only contract fails to achieve full supply chain efficiency. In the most common setting of a single supply chain wherein a supplier makes a contract offer to a retailer, the supplier can coordinate the supply chain by inducing the retailer to make decisions that are optimal to the supply chain as a whole. Such contracts may take a nonlinear form (e.g., quantity discount contracts) or include a fixed–payment component in addition to the wholesale price (e.g., two-part tariff contracts).

Although contract design has been extensively studied in the supply chain contracting literature, supply chain members' contractual form preferences have not. When supply chain members are profit maximizers, they should prefer a contractual form that yields higher profits. As a consequence, a profit-seeking supplier should prefer more sophisticated coordinating contracts to simple wholesale price–only contracts; with the former, the supplier can extract the entire supply chain surplus and leave none to the retailer. The retailer, however, should prefer the wholesale price–only contract because it allows him to secure a positive surplus (a result of the double marginalization effect). Thus, existing studies suggest that the supplier and retailer's contractual form preferences are not aligned and as a consequence there exist conflicts of interests among supply chain members with respect to supply chain contract design.

This paper focuses on the *congruence* of members' contractual form preferences. In particular, we focus on a market environment with deterministic demand and study the circumstances in which suppliers and retailers share the same preferences for a certain type of contract. Our analytical results suggest that both suppliers and retailers can be simultaneously better off using wholesale price–only contracts than other contractual forms. Therefore, the supplier and retailer's preferences can be *coordinated* on wholesale price–only contracts. There are also other situations in which both suppliers and retailers may be better using more complex contracts. We characterize conditions under which supply chain members share the same preferences for a certain type of contract, either wholesale price–only contract or quantity discount contract.

Supply chain structure can give members different degrees of market power. In order to understand how supply chain members' preferences change across different supply chain settings, we consider three forms of supply chain structure, all of which are selling same number of multiple products but differ in the number of suppliers or retailers as follows: (i) an $n$-supplier and $n$-retailer supply chain in which each supplier has an exclusive retailer; (ii) an $n$-supplier and 1-retailer supply chain in which all suppliers sell through a common retailer; and (iii) a 1-supplier and $n$-retailer supply chain in which one supplier sells through multiple retailers. The supply chain members in the three settings hold different levels of market power that is affected by the horizontal competition of multiple members.

We start by analyzing the optimal decisions at the retailer level. We then analyze the equilibrium decisions at the supplier level for a given type of contract while taking into account the retailers' responses. We choose two types of contracts for our analysis: the wholesale price–only contract and the quantity discount contract. The latter has been proved more efficient and complex by studies of a single supply chain. Our analysis shows that if all of the supply chains are restricted to a common contractual form (i.e., either a quantity discount contract or a wholesale price–only contract), then suppliers and retailers may both prefer the wholesale price–only contract or the quantity discount contract *simultaneously*, provided the competition intensity (as measured by the product substitution rate) falls within a certain range and the number of supply chains is large. The structure of the supply chain network plays a crucial role in determining which type of contract is preferred by the supply chain members.

This paper makes the following contributions. First, we develop an industry equilibrium analysis for competing supply chains under two commonly used contractual forms in three supply chain settings. More importantly, we show that, unlike the contracting in a 1-supplier and 1-retailer supply chain, suppliers and retailers may simultaneously prefer the wholesale price–only contract to more complex and efficient contracts and vice versa depending on the supply chain structure. Thus, our study offers a theoretical explanation to a long-standing dilemma in the literature that the wholesale price–only contract is a popular mechanism used in many supply chain transactions in practice, even though it has been extensively shown that in theory such contracts are clearly dominated by coordinating contracts. The congruence of such preferences may occur for a wide range of parameters when the number of competing supply chains is large. In this study, then, key industrial structure characteristics (e.g., structure of supply chain, product substitution, number of horizontal competitors) have crucial impacts on the congruence of preferences and the resulting prevalence of certain contractual forms.

## 4.2 Literature Review

Coordination by incentive contracts has been one of the core issues in the area of supply chain management because maximal profit for an entire supply chain cannot be achieved by commonly used wholesale price–only contracts in a simple supplier–retailer dyad (see Cachon 2003 and Nagarajan and Sosic 2008 for a comprehensive review of the literature. An exception that the wholesale price–only contract coordinates in the infinitely repeated interactions is shown by Sun and Debo 2014 ). Extensive study has been devoted to a variety of coordinating contracts that can align supply chain members' local incentives and thereby maximize supply chain profits. Examples include quantity discount (Cachon 2003, Tomlin 2003), quantity flexibility (Tsay 1999), sales rebate (Taylor 2002, Krishnan *et al.* 2004), two-part tariff (Cachon and Kök 2010), buy-back (Pasternack 1985), and revenue-sharing contracts (Cachon and Lariviere 2005a, Kannan and Popiuc 2014).

Although studies have not focused on supply chain members' contractual form preferences, existing research clearly suggests *uncoordinated* preferences for contractual forms of the upstream supplier and the downstream retailer: whereas the supplier should prefer coordinating contracts to wholesale price–only contracts, the retailer should prefer wholesale price-only contracts (to avoid being left with zero surplus, as may occur under coordinating contracts). Therefore, the literature on supply chain contracting implies an outstanding issue of conflict of interests between the supplier and the retailer in a supply chain. Furthermore, the wholesale price–only contract, a theoretically suboptimal contract for profit-maximizing suppliers, has been observed to govern supply chain transactions in many industries (Lariviere and Porteus 2001).

A recent study by Cachon and Kök [2010] sheds some light on the congruence of contractual form preferences. These authors consider a supply chain structure with two manufacturers selling substitutable products through a common retailer and show that, for intermediate levels of product substitution, manufacturers and retailers share the same preferences for more complex contracts (e.g., quantity discount and two-part tariff contracts) over wholesale price–only contracts. In a two-supplier and two-retailer supply chain setting, Feng and Lu [2013] study performance of supply chain contracting game that can be structured both as a Nash bargaining game and a Stackelberg game. They show that the way the contracting game is setup significantly influence supply chain members' contract choice.

In this paper, we examine different supply chain settings that involve multiple supply chain members engaging in horizontal competition at both the retailer and supplier levels. We find that when the substitution rate is high, suppliers and retailers both prefer wholesale price–only contracts to quantity discount contracts in an $n$-supplier and $n$-retailer setting. Preferences for contractual forms, however, are reversed when there are $n$ suppliers and 1 retailer: both the suppliers and the retailer prefer quantity discount contracts to wholesale price–only contracts. This indicates that supply chain structures affect the congruence of supply chain members' preferences for contractual forms.

Our work is also related to studies on the structure of supply chain channels beginning with the seminal paper of McGuire and Staelin [1983]. Following their basic model of two-stage competition, we extend the analysis to multiple supply chains and focus on the preferences for contractual forms in supply chain channels. In addition to analyzing the game with symmetric players, we also analyze the game with differentiated production costs among suppliers.

Finally, this paper is part of a growing field of research on supply chains with structures that are more complex than a simple supply chain dyad. In addition to the papers already cited, Bernstein and Federgruen [2005] and Bernstein *et al.* [2006] examine coordination in one-manufacturer and multiple-retailer supply chains with contracts. Netessine and Zhang [2005] consider both positive and negative externalities among downstream retailers and their effect on supply chain performance. Corbett and Karmarkar [2001] analyze horizontal competition in multi-stage supply chains with entry decisions. DeMiguel and Xu [2009] develop an equilibrium analysis of a multiple-leader and multiple-follower oligopoly game with stochastic demand. Adida and DeMiguel [2011] consider supply chain competition with uncertain demands and risk-averse decision makers. Federgruen and Hu [2012] study sequential oligopolies in supply chains featuring multiple echelons and firms that engage in price competition with other firms of the same echelon as well as in vertical competition across echelons.

In a setting of competing supply chains, Ha and Tong [2008] investigate suppliers' investment decisions on information sharing under different contract types. When the supplier's production exhibits diseconomy of scale in competitive supply chains, Ha *et al.* [2011] further study information sharing and its impact on supply chain members' performances . While our paper also considers both horizontal and vertical competition among multiple supply chain members, our focus is rather different. We are interested in when and how the preferences of the upstream supply side for contractual forms are (mis)aligned with those of the downstream retail side.

## 4.3 Model and Equilibrium Analysis

### 4.3.1 Supply Chain Structure

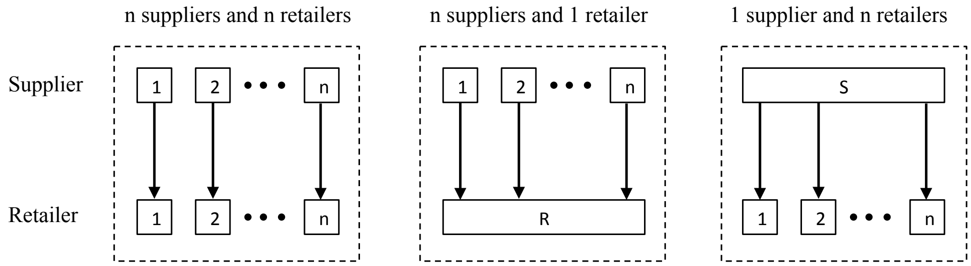
We consider three supply chain structures of $n$ partially substitutable products sold to a common market in which: (1) there are $n$ suppliers, each of whom is assumed to produce only one product and have one exclusive retailer; (2) $n$ suppliers sell their products via one common retailer; and (3) a single supplier produces $n$ products, each of which is sold via an exclusive retailer. Figure 1 illustrates the supply chain structures of this study. While in the $n$-supplier and 1-retailer and the 1-supplier and $n$-retailer supply chains horizontal competition takes place, respectively, at the supplier and retailer levels only, in the $n$-supplier and $n$-retailer supply chain, there are horizontal competitions among the retailers and among the suppliers (the competition among the suppliers , however, is indirect because their decisions affect each other only through the their retailer's decisions). In all three supply chain structures, the supplier is the Stackelberg leader and offers a contract to the retailer, taking into consideration the response function of the retailer.[1]

Product $i$'s inverse demand function is modeled as

$$p_i = \alpha_i - \beta d_i - \lambda \sum_{j \neq i} d_j \quad \text{for} \quad \beta > \lambda > 0 \text{ and } i = 1, 2, \ldots, n, \tag{4.1}$$

[1]While the supply chain structures in this study differ in the number of suppliers and retailers, they reflect different degrees of market power that the suppliers have. A supplier is the most powerful when selling through multiple competing retailers and facing no competition from other suppliers (i.e., in the 1-supplier and $n$-retailer supply chain) and the suppliers are the least powerful when competing with other suppliers to sell through a common retailer (i.e., in the $n$-supplier and 1-retailer supply chain). The suppliers in the $n$-supplier and $n$-retailer supply chain are more powerful than those in the $n$-supplier and 1-retailer supply chain because the horizontal competition at the retailer level alleviates the vertical competition between the suppliers and the retailers. Here, by comparing results across three different supply chain settings, our analysis illustrates the effects of the market power that is embedded in a particular supply chain structure on supply chain members' preferences for contractual forms.

Figure 4.1: Supply Chain Structures




where $p_i$ is the retail price of product $i$ and $d_i$ is the sales volume or equivalently the demand for that product. The parameter $\lambda$ represents the degree of product substitutability, which ranges from 0 for independent products to $\beta$ for perfect substitutes.

Because the suppliers form an oligopoly when there are $n$ suppliers in the $n$-supplier and $n$-retailer and the $n$-supplier and 1-retailer supply chain structures, each supplier also takes other suppliers' decisions into consideration when making her own contract offer. The retailers in the $n$-supplier and $n$-retailer and the 1-supplier and $n$-retailer structures also form an oligopoly, and they engage in Cournot competition.

In our model, the retailer's decision on retail prices is mathematically equivalent to choosing the sales volume $d_i$. Because the inverse demand functions in (4.1) are linear, product $i$'s demand $d_i$ can be solved as a function of $p_i$ as follows:

$$d_i = \frac{(\beta + \lambda(n-2))\,(\alpha_i - p_i) - \lambda \sum_{j \neq i}(\alpha_j - p_j)}{(\beta - \lambda)\,(\beta + \lambda(n-1))} \quad \text{for} \quad i = 1, 2, \ldots, n. \tag{4.2}$$

Let $t_i(d_i)$ be the transfer payment made by a retailer to supplier $i$ or to the common supplier based on the sales volume of product $i$. Product $i$ generates a profit of $p_i d_i - t_i(d_i)$ to the retailer, either to retailer $i$ of $n$ retailers or to the common retailer.

In the three supply chain structures, each supply chain member's profit is given as follows.

1. In the $n$-supplier and $n$-retailer supply chain, supplier $i$ and retailer $i$'s profits are

$$\pi_i^S = t_i(d_i) - c_i d_i, \quad \pi_i^R = p_i d_i - t_i(d_i). \tag{4.3}$$

2. In the $n$-supplier and 1-retailer supply chain, supplier $i$ and the retailer's profits are

$$\hat{\pi}_i^S = t_i(d_i) - c_i d_i, \quad \hat{\pi}^R = \sum_{i=1}^{n} \big(p_i d_i - t_i(d_i)\big). \tag{4.4}$$

3. In the 1-supplier and $n$-retailer supply chain, the supplier and retailer $i$'s profits are
$$\tilde{\pi}^S = \sum_{i=1}^{n} \big(t_i(d_i) - c_i d_i\big), \quad \tilde{\pi}_i^R = p_i d_i - t_i(d_i). \tag{4.5}$$

Here, $c_i \geq 0$ is the unit production cost of supplier $i$. We assume that the suppliers' and retailers' objectives are to maximize their profits shown in (4.3) - (4.5), respectively.

### 4.3.2 Contractual Forms

We consider a family of quantity discount contracts that the supplier can offer to the retailer. The functional form is the same as that used by Cachon and Kök [2010] and is fully characterized by parameters $w_i$ and $\upsilon_i$:
$$t_i(d_i) = \begin{cases} w_i d_i - \upsilon_i d_i^2/2, & \text{if } d_i \leq (w_i - c_i)/\upsilon_i, \\ t_i((w_i - c_i)/\upsilon_i) + c_i(d_i - (w_i - c_i)/\upsilon_i) & \text{otherwise.} \end{cases} \tag{4.6}$$
Here, $w_i$ is the wholesale price charged by supplier $i$, and the discount takes a quadratic form with discount rate $\upsilon_i \in [0, \bar{\upsilon})$, which is assumed to be less than $\bar{\upsilon} = 2\beta - \lambda$ in the $n$-supplier and $n$-retailer supply chain and in the 1-supplier and $n$-retailer supply chain, and to be less than $\bar{\upsilon} = 2\beta - 2\lambda$ in the $n$-supplier and 1-retailer supply chain. This is not a restrictive assumption because in our equilibrium analysis, retailer $i$'s quantity decision $d_i$ increases in his own wholesale price $w_i$ for $\upsilon_i > \bar{\upsilon}$. It is also intuitively true that the supplier incurs a negative profit if the discount rate is set too high. The wholesale price–only contract is a special case of the quantity discount contract: if $\upsilon_i = 0$, then $t_i(d_i) = w_i d_i$.

We focus on a positive discount rate contract ($\upsilon_i > 0$) and a wholesale price–only contract ($\upsilon_i = 0$). Therefore, we treat the discount rate $\upsilon$ as a parameter that denotes the contract type rather than as a decision variable. Because the wholesale price–only contract is a special case, we present the general analysis only for the quantity discount contract. The analysis of the wholesale price–only contract then follows directly by letting $\upsilon_i = 0$.

### 4.3.3 Equilibrium Analysis

#### 4.3.3.1 Supply Chains with $n$ Retailers

When there are $n$ retailers (either in the $n$-supplier and $n$-retailer setting or in the 1-supplier and $n$-retailer setting), retailer $i$'s best response function to the other $n-1$ retailers' sales volume decisions, conditional on the wholesale price $w_i$, is

$$d_i(d_{-i}) = \frac{\alpha_i - w_i - \lambda \sum_{j \neq i} d_j}{2\beta - \upsilon_i}, \quad \text{for } i = 1, 2, \ldots, n,$$

where $d_{-i} = (d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_n)'$ are the sales volumes of the other retailers. This system of equations has a unique solution and is given by

$$d_i(\mathbf{w}) \quad = \quad \frac{1 + \sum_{k \neq i}^{n} \frac{\varphi_k}{1-\varphi_k}}{\lambda \left(\frac{1-\varphi_i}{\varphi_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\varphi_k}{1-\varphi_k}\right)}(\alpha_i - w_i) - \frac{\sum_{j \neq i} \frac{\varphi_j}{1-\varphi_j}(\alpha_j - w_j)}{\lambda \left(\frac{1-\varphi_i}{\varphi_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\varphi_k}{1-\varphi_k}\right)}, \qquad (4.7)$$

where $\mathbf{w} = (w_1, w_2, \ldots, w_n)'$ and $\varphi_i = \frac{\lambda}{2\beta - \upsilon_i}$, for $i = 1, 2, \ldots, n$.

In the $n$-supplier and $n$-retailer supply chain, supplier $i$'s profit under the retailer's equilibrium sales (4.7) is given by

$$\pi_i^S(\mathbf{w}) = \left(w_i - \frac{\upsilon_i}{2} d_i(\mathbf{w}) - c_i\right) d_i(\mathbf{w}). \qquad (4.8)$$

In the 1-supplier and $n$-retailer supply chain, the supplier's profit under the retailer's equilibrium sales (4.7) is given by

$$\tilde{\pi}^S(\mathbf{w}) = \sum_{i=1}^{n} \left(w_i - \frac{\upsilon_i}{2} d_i(\mathbf{w}) - c_i\right) d_i(\mathbf{w}). \qquad (4.9)$$

#### 4.3.3.2 Supply Chains with One Retailer

In the $n$-supplier and 1-retailer setting, the retailer's profit (4.4) is a concave function of vector $\mathbf{d} = (d_1, d_2, \ldots, d_n)'$ (i.e., its Hessian matrix is strictly diagonally dominant for $\upsilon \leq \bar{\upsilon} = 2\beta - 2\lambda$), and the optimal solution is

$$\hat{d}_i(\mathbf{w}) = \frac{1 + \sum_{k \neq i}^{n} \frac{\hat{\varphi}_k}{1-\hat{\varphi}_k}}{2\lambda \left(\frac{1-\hat{\varphi}_i}{\hat{\varphi}_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\hat{\varphi}_k}{1-\hat{\varphi}_k}\right)}(\alpha_i - w_i) - \frac{\sum_{j \neq i} \frac{\hat{\varphi}_j}{1-\hat{\varphi}_j}(\alpha_j - w_j)}{2\lambda \left(\frac{1-\hat{\varphi}_i}{\hat{\varphi}_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\hat{\varphi}_k}{1-\hat{\varphi}_k}\right)}, \qquad (4.10)$$

where $\hat{\varphi}_i = \frac{2\lambda}{2\beta - \upsilon_i}$, for $i = 1, 2, \ldots, n$.

Supplier $i$'s profit under the retailer's optimal sales volume (4.10) is given by

$$\hat{\pi}_i^S(\mathbf{w}) = \left(w_i - \frac{\upsilon_i}{2} \hat{d}_i(\mathbf{w}) - c_i\right) \hat{d}_i(\mathbf{w}). \qquad (4.11)$$

In addition to the uniqueness of the retailer sales decisions in all three supply chain structures, we also have the uniqueness of the equilibrium wholesale prices of the supplier competition. This result is stated formally in the following theorem.

**Theorem 14** *Equilibrium wholesale prices exist and are unique in all three supply chain structures.*

All proofs are given in the Appendix D. Because the selling quantities are uniquely determined at the retailer level, Theorem 14 implies that the entire two-stage game has a unique equilibrium for any given discount rate in each of the three supply chain structures.

## 4.4 Preferences for Contractual Forms

### 4.4.1 Base Case

We first analyze the base case of a one-supplier–one-retailer dyad. Note that a single supply chain dyad is a special case of our competing supply chain models with no product substitution (i.e., $\lambda = 0$). Conditional on a contract offer with a discount rate $\upsilon$, the retailer's best-response function is $d = \frac{\alpha - w}{2\beta - \upsilon}$. Substituting this into the supplier's profit function, we obtain the optimal wholesale price offer $w^* = \frac{2\beta\alpha + (2\beta - \upsilon)c}{4\beta - \upsilon}$ and thus the equilibrium sales volume $d^* = \frac{\alpha - c}{4\beta - \upsilon}$. In equilibrium, the supplier earns $\pi^S = \frac{(\alpha - c)^2}{8\beta - 2\upsilon}$ while the retailer earns $\pi^R = \frac{(2\beta - \upsilon)(\alpha - c)^2}{2(4\beta - \upsilon)^2}$.

In terms of contractual forms, the supplier always prefers a quantity discount contract because $\frac{\partial \pi^S}{\partial \upsilon} > 0$, and the retailer always prefers a wholesale price–only contract because $\frac{\partial \pi^R}{\partial \upsilon} < 0$. As described here, supply chain members' preferences for contractual forms in a single supply chain are consistent with the implications of the supply chain coordination literature. From the supplier's perspective, the quantity discount contract is always preferable. However, the wholesale price–only contract is preferable from the retailer's perspective because it yields him the surplus resulting from double marginalization. Therefore, in the single supply chain setting, the upstream and downstream parties inevitably have conflicting preferences.

### 4.4.2 Supply Chains with Upstream or Downstream Competition

We now extend our analysis to allow competition among the multiple supply chain members modelled in the previous section (i.e., $\lambda > 0$). To focus on contractual form preferences, we assume that the $n$ products are symmetric in terms of market size and production cost; that is, $\alpha_i - c_i = \xi$ for all $i = 1, 2, \ldots, n$. Even so, the products may differ in both $\alpha_i$ and $c_i$. The parameter $\alpha_i$ in the inverse demand function is equivalent to the most a customer is willing to pay for product $i$; hence the term $\alpha_i - c_i$ represents the maximum profit margin that product $i$ can generate. By assuming constant $\alpha_i - c_i$ across products, profit differences under different contracts are driven only by the contractual type.

Since our focus is on the contractual form preferences, we investigate the case in which all supply chains use the same type of contract; thus, $\upsilon_i = \upsilon$, where $\upsilon = 0$ represents the simple wholesale price–only contract and $\upsilon > 0$ the more complex quantity discount contract.[2] This is not an unusual setting and has been observed in practice. Cachon and Kök [2010] cite "anecdotal evidence from our conversations with executives from grocery retail chains (H-E-B in Texas and Stop & Shop in the northeastern United States) that, in most categories, most firms offer the same type of contracts." The following lemma characterizes the equilibrium of the entire game.

**Lemma 5** (Equilibrium Wholesale Prices and Sales Volumes )
*Let $\upsilon_i = \upsilon$ and $\alpha_i - c_i = \xi$ for $i = 1, 2, \ldots, n$. The equilibrium wholesale prices and sales volumes are then given as follows:*

1. *For the $n$-supplier and $n$-retailer supply chain, let $\varphi = \lambda/(2\beta - \upsilon)$, then*

$$\begin{aligned} w_i^* &= \alpha_i - \frac{(2\beta-\upsilon)[1+(n-2)\varphi][1+(n-1)\varphi]\xi}{2\beta[2+(n-3)\varphi][1+(n-1)\varphi] - \upsilon\big[1+\varphi[2n-3+(n-3)(n-1)\varphi]\big]}, \qquad (4.12)\\ d_i^* &= \frac{[1+(n-2)\varphi]\xi}{2\beta[2+(n-3)\varphi][1+(n-1)\varphi] - \upsilon\big[1+\varphi[2n-3+(n-3)(n-1)\varphi]\big]}. \end{aligned}$$

2. *For the $n$-supplier and 1-retailer supply chain, let $\hat{\varphi} = 2\lambda/(2\beta - \upsilon)$, then*

$$\begin{aligned} \hat{w}_i^* &= \alpha_i - \frac{(2\beta-\upsilon)[1+(n-2)\hat{\varphi}][1+(n-1)\hat{\varphi}]\xi}{2\beta[2+(n-3)\hat{\varphi}][1+(n-1)\hat{\varphi}] - \upsilon\big[1+\hat{\varphi}[2n-3+(n-3)(n-1)\hat{\varphi}]\big]},\\ \hat{d}_i^* &= \frac{[1+(n-2)\hat{\varphi}]\xi}{2\beta[2+(n-3)\hat{\varphi}][1+(n-1)\hat{\varphi}] - \upsilon\big[1+\hat{\varphi}[2n-3+(n-3)(n-1)\hat{\varphi}]\big]}. \end{aligned}$$

3. *For the 1-supplier and $n$-retailer supply chain, let $\varphi = \lambda/(2\beta - \upsilon)$, then*

$$\begin{aligned} \tilde{w}_i^* &= \alpha_i - \frac{\lambda(1+(n-1)\varphi)}{2\lambda(1+(n-1)\varphi)+\upsilon\varphi}\xi,\\ \tilde{d}_i^* &= \frac{\varphi}{2\lambda(1+(-1+n)\varphi)+\upsilon\varphi}\xi. \end{aligned}$$

[2] While we assume that all suppliers are restricted to use either a quantity discount contract or a wholesale price–only contract, it is possible to allow each supplier to choose the type of contract, in particular in the $n$-supplier and $n$-retailer setting when each supplier sells through an exclusive retailer. The analysis of this case can be found in Lu and Wu [2013].

In equilibrium, the supplier and retailer's profits are given (respectively) as follows.[3]

1. For the $n$-supplier and $n$-retailer supply chain, let $\varphi = \lambda/(2\beta - \upsilon)$,

$$\pi_i^S(\upsilon) = \frac{4\beta(1-\varphi)[1+(n-2)\varphi][1+(n-1)\varphi] - \upsilon[1+(n-2)\varphi]\big[1+(n-2)\varphi - 2(n-1)\varphi^2\big]}{2\Big[2\beta[2+(n-3)\varphi][1+(n-1)\varphi] - \upsilon\big(1+\varphi[2n-3+(n-3)(n-1)\varphi]\big)\Big]^2}\xi^2,$$

$$\pi_i^R(\upsilon) = \frac{(2\beta-\upsilon)[1+(n-2)\varphi]^2}{2\Big[2\beta[2+(n-3)\varphi][1+(n-1)\varphi] - \upsilon\big(1+\varphi[2n-3+(n-3)(n-1)\varphi]\big)\Big]^2}\xi^2.$$

2. For the $n$-supplier and 1-retailer supply chain, let $\hat{\varphi} = 2\lambda/(2\beta - \upsilon)$,

$$\hat{\pi}_i^S(\upsilon) = \frac{4\beta(1-\hat{\varphi})[1+(n-2)\hat{\varphi}][1+(n-1)\hat{\varphi}] - \upsilon[1+(n-2)\hat{\varphi}]\big[1+(n-2)\hat{\varphi} - 2(n-1)\hat{\varphi}^2\big]}{2\Big[2\beta[2+(n-3)\hat{\varphi}][1+(n-1)\hat{\varphi}] - \upsilon\big(1+\hat{\varphi}[2n-3+(n-3)(n-1)\hat{\varphi}]\big)\Big]^2}\xi^2,$$

$$\hat{\pi}^R(\upsilon) = \frac{n(2\beta-\upsilon)[1+(n-2)\hat{\varphi}]^2[1+(n-1)\varphi]}{2\Big[2\beta[2+(n-3)\hat{\varphi}][1+(n-1)\hat{\varphi}] - \upsilon\big(1+\hat{\varphi}[2n-3+(n-3)(n-1)\hat{\varphi}]\big)\Big]^2}\xi^2.$$

3. For the 1-supplier and $n$-retailer supply chain, let $\varphi = \lambda/(2\beta - \upsilon)$,

$$\tilde{\pi}^S = \frac{n\varphi}{4\lambda(1+(n-1)\varphi) + 2\upsilon\varphi}\xi^2,$$

$$\tilde{\pi}_i^R = \frac{\varphi[2\lambda - \varphi(2\beta-\upsilon)]}{2[2\lambda(1+(n-1)\varphi + \varphi\upsilon)]^2}\xi^2.$$

To study contractual form preferences under a given industrial structure, our analysis focuses mainly on conditions under which the supplier can gain a higher profit by offering a wholesale price-only contract, and under which the retailer can gain a higher profit by taking a wholesale price-only contract. The following thorem characterizes these conditions.

**Theorem 15** (PREFERENCES OF SUPPLY CHAIN MEMBERS)
*If all of the supply chain members take either the wholesale price–only contract ($\upsilon = 0$) or the quantity discount contract ($\upsilon > 0$), their contractual form preferences are determined as follows.*

1. *For the $n$-supplier and $n$-retailer supply chain, there exist $\lambda^S(n), \lambda^R(n) \in (0, \beta]$ such that the suppliers prefer the wholesale price–only contract when $\lambda \geq \lambda^S(n)$ and the retailers prefer the wholesale price–only contract when $\lambda \leq \lambda^R(n)$. Furthermore, $\lambda^R(n) = \beta$ and $\lambda^R(n) > \lambda^S(n)$ for any $n \geq 3$.*

[3]In our analysis, it is true that the equilibrium market price is nonnegative and all supply chain members earn nonnegative profits. Details are shown in the Appendix.

2. *For the n-supplier and 1-retailer supply chain, there exist* $\hat{\lambda}^S(n), \hat{\lambda}^R(n) \in (0, \beta - \upsilon/2]$ *such that the suppliers prefer the wholesale price–only contract when* $\lambda \geq \hat{\lambda}^S(n)$ *and the retailers prefer the wholesale price–only contract when* $\lambda \leq \hat{\lambda}^R(n)$*. Furthermore,* $\hat{\lambda}^R(n) < \hat{\lambda}^S(n)$ *for any* $n \geq 2$*.*

3. *For the 1-supplier and n-retailer supply chain, the supplier always prefer the quantity-discount contract and the retailers always prefer the wholesale price–only contract.*

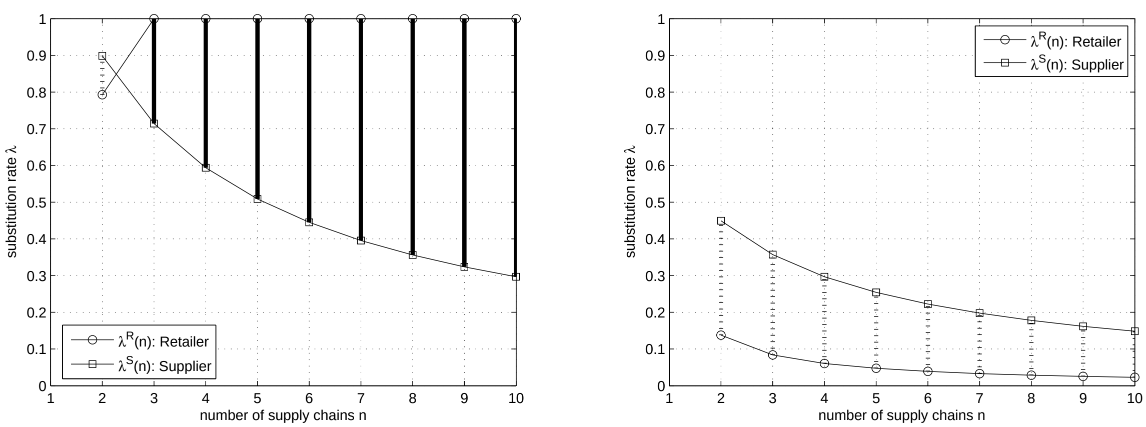
This theorem shows that the supply chain structure along with the competition intensity captured by product substitution play key roles in the contractual preferences of supply chain members. Members of the 1-supplier and $n$-retailer supply chain maintain the same preferences as in a single supply chain dyad. Because the horizontal competition in this setting only occurs at the retailer level, this result implies that the competition among suppliers is an important driver of suppliers' contractual form preferences.

In the following analysis, we focus on the other two supply chain settings that have $n$ suppliers. In those two settings, suppliers and retailers' preferences are both determined mainly by two forces of competition: the product substitution rate $\lambda$ and the level of congestion $n$ in the market. Suppliers prefer the wholesale price–only contract when the substitution rate is high because the competition among suppliers alone reduces wholesale price and thus increases order quantities, in which case it becomes less attractive for suppliers to offer a discount. However, when the product substitution rate is low, supply chains become more independent. In this case, offering quantity-based discounts provides an incentive for the retailer to order more and benefit the supplier. Retailers, in contrast, prefer the wholesale price–only contract when the product substitution rate is low because it allows them to increase profits owing to double marginalization.

Furthermore, we find that in the $n$-supplier and 1-retailer setting, the threshold of the supplier is always higher than that of the retailer, i.e., $\hat{\lambda}^S(n) > \hat{\lambda}^R(n)$. This result suggests that in such a setting, the congruence of preferences can only occur on the quantity discount contract. When there are $n$ suppliers and $n$ retailers, the thresholds are such that $\lambda^S(n) < \lambda^R(n)$ for $n > 3$, suggesting the congruence of preferences occurs only on the wholesale price–only contract. When $n = 2$ in the n-supplier and n-retailer setting, the congruence of preferences can occur either on the wholesale price–only contract or on the quantity discount contract. We report more detailed results for $n = 2$ in Section 4.5.

Figure 4.2(a) illustrates a preference congruence for the $n$-supplier and $n$-retailer setting with the following parameter values: $\alpha = \beta = 1$, $c = 0.05$, and $\upsilon = 0.8$. If $\lambda^S(n) < \lambda^R(n)$ (for $n \geq 3$), then suppliers and retailers both prefer the wholesale price–only contract when the substitution rate $\lambda$ satisfies $\lambda^S(n) \leq \lambda \leq \lambda^R(n)$; these cases are indicated by the dark vertical lines in the figure. If $\lambda^S(n) > \lambda^R(n)$ (i.e., $n = 2$), then suppliers and retailers both prefer the quantity discount contract when $\lambda^R(n) \leq \lambda \leq \lambda^S(n)$, as indicated by the "dotted" vertical lines. In Figure 4.2(b), we illustrate the preference congruence of the $n$-supplier and 1-retailer setting with the same set of parameters. Here, the congruence occurs on the quantity discount contract, indicated by the "dotted" vertical lines.

Figure 4.2: Preferences for the supply chain members




(a) $n$-supplier and $n$-retailer

(b) $n$-supplier and 1-retailer

In comparing these two settings, surprisingly, we observe preference reversals caused by the number of retailers: the preference congruence takes place on the wholesale price–only contract when there are multiple retailers, and on the quantity discount contract when there is a common retailer. Since the only difference between these two supply chain structures is the existence of horizontal competition between retailers, this competition drives the common preferences for the wholesale price–only contract. In particular, the competition among the retailers reduce each other's profit compared with the cases of independent retailers. As a consequence, the retailers prefer the wholesale price–only contract more often under which the double marginalization effect leaves more profit to them than under the quantity discount contract when multiple retailers compete at the retail market (Comparing Figures 2.(a) and 2.(b), we find that the threshold values below which the retailers prefer wholesale price–only contract are lower in the $n$-supplier and 1-retailer supply chain).

As discussed previously, studies on single supply chains indicate that the supplier always prefers the (more efficient) quantity discount contract, whereas the retailer always prefers the (simpler) wholesale price–only contract. As a result, the supplier and retailer *always* conflict in their preferences for contractual forms. However, in the setting of multiple competing supply chains, we show that contractual form preferences are not universal over varying degrees of product substitution. Even more interesting is that a congruence of preferences among different supply chain members can occur as a result of the underlying competing supply chain structure.

We briefly discuss how the profit of the entire supply chain system is affected by different types of contract. In the following Figure 3, we illustrate that relative to the wholesale price–only contract, how the improvement of supply chain profit by using the quantity discount contract depends on the substitution rate $\lambda$ in the $n$-supplier and $n$-retailer setting. Similar to the supplier's preference, the supply chain profit can be improved by using the wholesale price when the product substation rate is high. In other words, the entire supply chain system benefits from using the wholesale price–only contact when $\lambda$ is large. The reason is both sides of the supply chain prefer the wholesale price–only contract for large $\lambda$ as stated in Theorem 2. In the other two settings, there is no congruence of preferences on the wholesale price–only contract, and the quantity discount contract always improves supply chain profits.

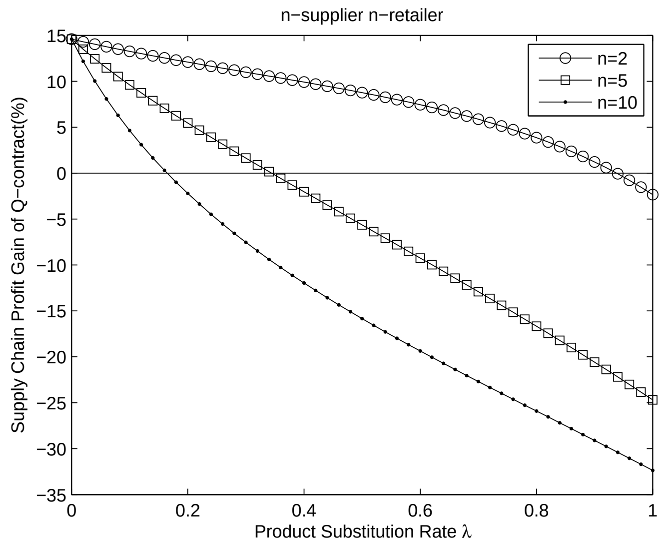
### 4.4.3 Preferences under Large Number of Products

Figure 4.2 also shows that the thresholds of the supplier $\lambda^S(n)$ and $\hat{\lambda}^S(n)$ are decreasing in $n$, which means that the suppliers in these two settings prefer simple wholesale price–only contracts more often when there are more competing supply chains. The following theorem formally states the preferences when the number of products is large.

**Theorem 16** (Preference under Large Number of Products )

1. *For the $n$-supplier and $n$-retailer supply chain, there exists an $N_0 \in [2, \infty)$ such that suppliers and retailers both prefer the wholesale price–only contract for any $\lambda > 0$ when $n \geq N_0$.*

Figure 4.3: Supply chain profit gain in the $n$-supplier and $n$-retailer setting




*Notes:* The parameters for both figures are: $\alpha = \beta = 1, c = 0.05, v = 0.8$. Supply chain profit gain by quantity discount contract is defined as follows: $\frac{\Pi_{SC}(QD)-\Pi_{SC}(W)}{\Pi_{SC}(W)} \times 100\%$, where $\Pi_{SC}(QD)$ and $\Pi_{SC}(W)$, respectively, represent supply chain profits by using the quantity discount contract and by using the wholesale price contract.

2. *For the n-supplier and 1-retailer supply chain, there exists an $\hat{N}_0 \in [2, \infty)$ such that suppliers prefer the wholesale price–only contract for any $\lambda > 0$ when $n \geq \hat{N}_0$.*

In terms of the retailer preference in the $n$-supplier and 1-retailer case, our analysis shows that when $n$ goes to infinity, the retailer is indifferent between the wholesale price–only and quantity discount contracts. Our numerical analysis reveals that the threshold for the retailer $\hat{\lambda}^R(n)$ decreases in $n$, which suggests that the retailer prefers the quantity discount contract for large $n$. The preferences in the 1-supplier and $n$-retailer setting are the same as stated in Theorem 15.

This Theorem states that, in a market with a large number of suppliers, the suppliers prefer the wholesale price–only contract as long as their products are substitutable — no matter how low the rate is. This result formalizes the observations made in Figure 4.2. If we extend the number of supply chains in that figure, then the curve continues to decrease and approaches zero as $n$ becomes very large.

### 4.4.4 Preferences under Highest Discount $\upsilon$

We have analyzed the equilibrium preferences in Theorem 15 under intermediate levels of the discount rate (i.e., $0 < \upsilon < \bar{\upsilon}$). When this discount rate is at its highest possible value, the thresholds can be expressed in closed-form. The equilibrium profits for each supply chain member under the wholesale price–only contract ($\upsilon = 0$) and the quantity discount contract ($\upsilon \to \bar{\upsilon}$) are reported in Table 1. Theorem 17 further characterizes contractual form preferences when the discount rate reaches the upper limit of $\bar{\upsilon}$.

Table 4.1: Equilibrium profit under wholesale price–only contract and quantity discount contract

| | Contract form | |
|---|---|---|
| | Wholesale price–only | Quantity discount as $\upsilon \to \bar{\upsilon}$ |
| Supplier (n) | $\frac{(2\beta-\lambda)(2\beta+\lambda(n-2))}{(4\beta+\lambda(n-3))^2(2\beta+\lambda(n-1))}\xi^2$ | $\frac{2\beta-\lambda}{2(2\beta+\lambda(n-1))^2}\xi^2$ |
| Retailer (n) | $\frac{\beta(2\beta+\lambda(n-2))^2}{(4\beta+\lambda(n-3))^2(2\beta+\lambda(n-1))^2}\xi^2$ | $\frac{\lambda}{2(2\beta+\lambda(n-1))^2}\xi^2$ |
| Supplier (n) | $\frac{(\beta-\lambda)(\beta+\lambda(n-2))}{2(2\beta+\lambda(n-3))^2(\beta+\lambda(n-1))}\xi^2$ | $\frac{\beta-\lambda}{4(\beta+\lambda(n-1))^2}\xi^2$ |
| Retailer (1) | $\frac{n(\beta+\lambda(n-2))^2}{4(2\beta+\lambda(n-3))^2(\beta+\lambda(n-1))}\xi^2$ | $\frac{n^2\lambda}{4(\beta+\lambda(n-1))^2}\xi^2$ |
| Supplier (1) | $\frac{n}{8\beta+4\lambda(n-1)}\xi^2$ | $\frac{n}{4\beta+2\lambda(2n-1)}\xi^2$ |
| Retailer (n) | $\frac{\beta}{4[2\beta+\lambda(n-1)]^2}\xi^2$ | $\frac{\lambda}{2[2\beta+\lambda(2n-1)]^2}\xi^2$ |

*Notes:* (a) Supply chain structure is indicated by the supplier and retailer numbers in brackets. (b) For the $n$-supplier and $n$-retailer and 1-supplier and $n$-retailer cases, $\bar{\upsilon} = 2\beta - \lambda$. For the $n$-supplier and 1-retailer case, $\bar{\upsilon} = 2\beta - 2\lambda$.

**Theorem 17** (Preference with Highest Discount)
*If the discount rate is set at its highest possible level (i.e., $\upsilon \approx \bar{\upsilon}$), then the following statements hold.*

1. *For $\upsilon \approx \bar{\upsilon} = 2\beta - \lambda$ in the n-supplier and n-retailer supply chain, the suppliers prefer the wholesale price–only contract to the quantity discount contract if and only if $\lambda > \frac{-6+2\sqrt{2n^2-1}}{n^2-5}\beta$. The retailers prefer the wholesale price–only contract to the quantity discount contract for all values of $\lambda$ when $n \geq 3$. The retailers prefer the wholesale price–only contract if and only if $\lambda < (3-\sqrt{5})\beta$ when $n = 2$.*

2. *For $\upsilon \approx \bar{\upsilon} = 2\beta - 2\lambda$ in the n-supplier and 1-retailer supply chain, the suppliers prefer the wholesale price–only contract to the quantity discount contract if and only if $\lambda > \frac{-3+\sqrt{2n^2-1}}{n^2-5}\beta$. The retailer prefers the wholesale price–only contract to the quantity discount contract if and only if $\lambda < \frac{2}{4+n+\sqrt{n(4+5n)}}\beta$.*

3. *For $\upsilon \approx \bar{\upsilon} = 2\beta - \lambda$ in the 1-supplier and n-retailer supply chain, the supplier always prefers the quantity–discount contract. The retailers always prefers the wholesale price–only contract.*

These results confirm the insights obtained from Theorem 15. When there are $n$ suppliers, they prefer the wholesale price–only contract if the substitution rate is high ($\lambda > \frac{-6+2\sqrt{2n^2-1}}{n^2-5}\beta$ and $\lambda > \frac{-3+\sqrt{2n^2-1}}{n^2-5}\beta$ in cases 1 and 2 respectively); retailers prefer the wholesale price when the substitution rate is low ($\lambda < (3-\sqrt{5})\beta$ and $\lambda < \frac{2}{4+n+\sqrt{n(4+5n)}}\beta$ in cases 1 and 2 respectively) if there are two supply chains, and they always prefer it when facing more than two competitors.

In addition, the number of competing supply chains also has an effect similar to that of the substitution rate; in particular, the threshold values in Theorem 4 are all decreasing in $n$. Moreover, the closed-form threshold values for the substitution rate also provide qualitative insights into when a certain contractual form is prevalent.

### 4.4.5 A Special Case of $n = 2$

We further illustrate the effect of supply chain structure on preference congruence via numerical examples in the settings of $n = 2$ (we set $\beta = 1$ in all examples). Figure 4.4(a) shows a congruence of supply chain members in the setting of two suppliers and two retailers. If $\upsilon = 0.8$ (the upper panel), then $\lambda^S = 0.898$ and $\lambda^R = 0.793$. For $0.793 \leq \lambda \leq 0.898$, suppliers and retailers both prefer the quantity discount contract (indicated by the shaded area). When $\upsilon = 1.2$ (the middle panel), we find that $\lambda^S = 0.750$ and $\lambda^R = 0.770$; for any $\lambda$ that falls between these two thresholds, both suppliers and retailers prefer the wholesale price–only contract (again indicated by the shaded area). When the discount is given by the highest $\upsilon \approx 2\beta - \lambda$ (the lower panel), we find that the preferences of supply chain members are similar to those in the scenario of $\upsilon = 1.2$, with $\lambda^S = 0.708$ and $\lambda^R = 0.764$. Thus, the *congruence* of contractual form preferences occurs for intermediate degrees of product substitution, although the contract form that is actually preferred depends on the values of the other parameters.

Further, the preferences of supply chain members in the 2-supplier and 1-retailer setting are presented in Figure 4.4(b). Here, the congruence can only be on the more efficient contract, i.e., the quantity discount contract (as indicated in the shaded areas in all three figures). Therefore, we find that the competition intensity and supply chain structures play crucial roles in supply chain members' contractual choices.[4]

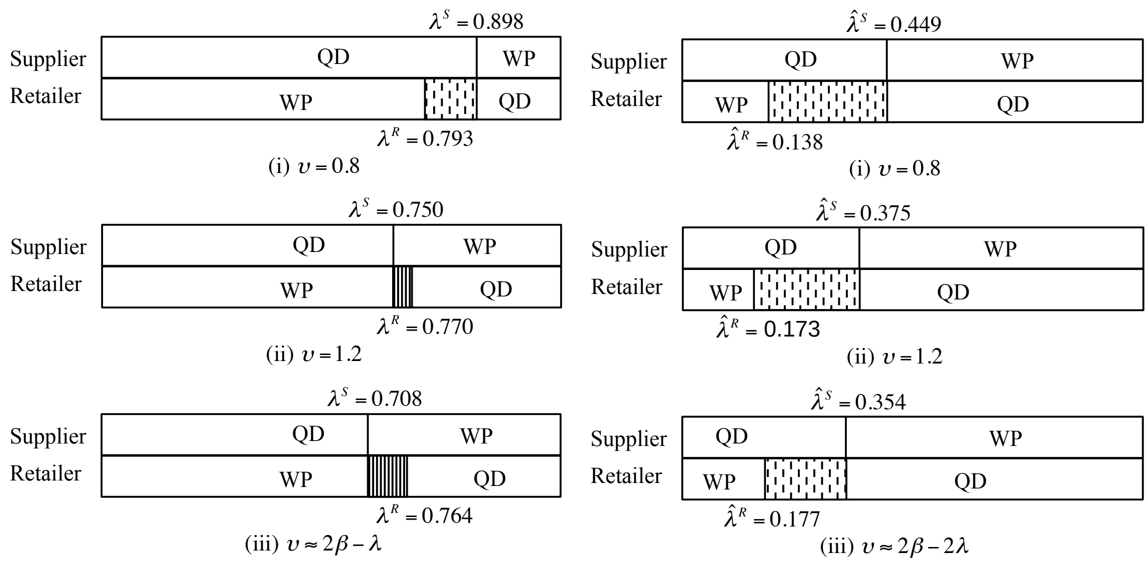
## 4.5 Conclusion

Contracts play a critical role in coordinating the activities of different parties in supply chains. Various contractual forms have been extensively examined in the supply chain management literature. It is a well-established theoretical result in the supply chain coordination literature that simple wholesale price–only contracts fail to achieve the full profit potential of a supply chain dyad when both the upstream supplier and the downstream retailer maximize their own profits. The upstream supplier is better off under the (more efficient) quantity discount contract whereas the downstream retailer is better off under the (less complex) wholesale price–only contract. Hence, the supply and demand sides of the chain are at odds with respect to preferred contractual forms.

In this paper, we focus on supply chain members' contractual form preferences in three different supply chain settings. We choose two contractual forms: the wholesale price–only contract and the quantity discount contract. We characterize the equilibria of the supply chain system and

[4] In their 2-supplier and 1-retailer setting, Cachon and Kök (2010) show that the suppliers prefer the wholesale price–only contract if and only if $\lambda \geq (3 - \sqrt{7})\beta \approx 0.354\beta$, and the retailer prefers the wholesale price–only contract if and only if $\lambda \leq \frac{3-\sqrt{7}}{2}\beta \approx 0.177\beta$, as shown in the lower panel of Figure 2.(b). Consequently, the preference congruence of the supply chain members only occurs such that both prefer the quantity discount contract. Feng and Lu (2013) show that all supply chain members may prefer the wholesale price-only contract in their 2-supplier and 2-retailer supply chain setting. Therefore, the supply chain structure has a crucial role in determining contract preferences among supply chain members.

Figure 4.4: Contract preference of supply chain members (WP - wholesale price only; QD - quantity discount)




(a) The 2-supplier and 2-retailer case (b) The 2-supplier and 1-retailer case

investigate the conditions under which supply chain members prefer a certain contractual form. Our analysis indicates that the competition among supply chains has a significant effect on each member's performance under different contractual forms. Suppliers and retailers can both be better off with wholesale price–only or quantity–discount contracts, which are systemically affected by the supply chain structure. The analysis presented here identifies conditions under which supply chain members' contractual form preferences are congruent, which, according to existing studies on the single supply chain, is not possible.

We use a parsimonious competition model to examine preference congruences in supply chains. Extending this model to more complex supply chain systems would test the robustness of our findings. The literature on supply chain contracting and coordination has shown how different contractual forms (e.g., two-part tariff, revenue sharing, buy-back, etc.) can achieve the same system efficiency. It would be interesting to study supply chain member's preferences when multiple contractual forms are available, or between the wholesale price–only contract and other more complex contractual forms. Future work can also examine contractual preferences under demand uncertainty, which would introduce another important factor that may have significant influence on the preferences for contracts. We hope that extending the analysis of misaligned interests in a single supply chain to the competition among multiple supply chains will generate more insights on the interaction between industrial structure and supply chain contracting.

# Bibliography

H. Aaron and A. Frakt. Why now is not the time for premium support. *New England Journal of Medicine*, 366:877–879, 2012.

E. Adida and V. DeMiguel. Supply chain competition with multiple manufacturers and retailers. *Operations Research*, 59:156–172, 2011.

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

S. Alizamir, F. de Vericourt, and P. Sun. Efficient feed-in-tariff policies for renewable energy technologies. To appear in Operations Research, 2016.

G. Allon, A. Federgruen, and M. Pierson. Price competition under multinomial logit demand functions with random coefficients. *Management Science*, 59:1817–1835, 2013.

S. Anderson, A. Palma, and J. Thisse. *Discrete choice theory of product differentiation*. The MIT Press, 2001.

Anonymous. A procurement auction model for framework agreements. In *Charting a Course in Public Procurement Innovation and Knowledge Sharing*. 2012 International Public Procurement Conference Publications, edited by PrAcademics Press, 2012.

J. Antos. The wyden-ryan proposal–a foundation for realistic medicare reform. *New England Journal of Medicien*, 366:879–881, 2012.

K. Arrow, H. Block, and L. Hurwicz. On the stability of competitive equilibrium, ii. *Econometrica*, 27:82–109, 1959.

S. Athey. Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69:861–889, 2001.

D. Bell and J. Lattin. Looking for loss aversion in scanner panel data: The confounding effect of price response heterogeneity. *Marketing Science*, 19:185–200, 2000.

F. Bernstein and A. Federgruen. A general equilibrium model for industries with price and service competition. *Operations Research*, 52:868–886, 2004.

F. Bernstein and A. Federgruen. Decentralized supply chains with competing retailers under demand uncertainty. *Management Science*, 51:18–29, 2005.

F. Bernstein, F. Chen, and A. Federgruen. Coordinating supply chains with simple pricing schemes: The role of vendor-managed inventories. *Management Science*, 52:1483–1492, 2006.

J. Bertrand. Review of Cournot's 'rechercher sur la theorie mathematique de la richesse'. *Journal des Savants*, pages 499–508, 1883.

D. Besanko, M.K. Perry, and R.H. Spady. The logit model of monopolistic competition: Brand diversity. *The Journal of Industrial Economics*, pages 397–415, 1990.

R. Briesch, L. Krishnamurthi, T. Mazumdar, and S. Raj. A comparative analysis of reference price models. *Journal of Consumer Research*, 24:202–214, 1997.

J. Brown, M. Duggan, I. Kuziemko, and W. Woolston. How does risk selection respond to risk adjustment? new evidence from the medicare advantage program. *American Economic Review*, 104:3335–3364, 2014.

L. Cabral and M. Villas-Boas. Bertrand supertraps. *Management Science*, 51:599–613, 2005.

M. Cabral, M. Geruso, and N. Mahoney. Does privatized medicare benefit patients or producers? evidence from the benefits improvement and protection act. NBER Working Paper No. 20470, 2014.

G. Cachon and A. Kök. Competing manufacturers in a retail supply chain: On contractual form and coordination. *Management Science*, 56:571–589, 2010.

G. Cachon and M. A. Lariviere. Supply chain coordination with revenue-sharing contracts: Strengths and limitations. *Management Science*, 51:30–44, 2005.

G. Cachon and M. A. Lariviere. Supply chain coordination with revenue-sharing contracts: Strengths and limitations. *Management Science*, 51:30–44, 2005.

G. Cachon. *Supply chain coordination with contracts.* A. G. de Kok, S. C. Graves, eds. *Handbooks in Operations Research and Management Science: Supply Chain Management*, Chap. 6. Elsevier, Amsterdam, 2003.

E. Cadotte, R. Woodruff, and R. Jenkins. Expectations and norms in models of consumer satisfaction. *Journal of Marketing Research*, 24:305–314, 1987.

L. Cahn. How to combat pharma's costly coupon programs. *Managed Care*, 21:22–26, 2012.

A. Caplin and B. Nalebuff. Aggregation and imperfect competition: On the existence of equilibrium. *Econometrica*, 59:25–59, 1991.

X. Chen, P. Hu, and Z. Hu. Efficient algorithms for dynamic pricing problem with reference price effect. To appear in Operations Research, 2016.

F. Chen. Auctioning supply contracts. *Management Science*, 53:1562–1576, 2007.

M. Cohen, R. Lobel, and G. Perakis. The impact of demand uncertainty on consumer subsidies for green technology adoption. To appear in Management Science, 2016.

Congressional Budget Office. A premium support system for medicare: analysis of illustrative options. http://www.cbo.gov/sites/default/files/cbofiles/attachments/09-18-PremiumSupport.pdf, 2013.

C. Corbett and U. Karmarkar. Competition and structure in serial supply chains with deterministic demand. *Management Science*, 47:966–978, 2001.

V. Curto, L. Einav, J. Levin, and J. Bhattacharya. Can health insurance competition work? evidence from medicare advantage. NBER Working Paper No. 20818, 2015.

V. DeMiguel and H. Xu. A stochastic multiple-leader stackelberg model: Analysis, computation, and application. *Operations Research*, 57:1220–1235, 2009.

B. Dowd, R. Feldman, and R. Coulam. The effect of health plan characteristics on medicare+choice enrollment. *Health Services Research*, 38:113–135, 2003.

J. Dube, G. Hitsch, and P. Rossi. Do switching costs make markets less competitive? *Journal of Marketing Research*, 46:435–445, 2009.

I. Duenyas, B. Hu, and D. Beil. Simple auctions for supply contracts. Forthcoming, Management Science, 2013.

M. Duggan, A. Starc, and B. Vabson. Who benefits when the government pays more? pass-through in the medicare advantage program. NBER Working Paper No. 19989, 2014.

L. Einav and J. Levin. Managed competition in health insurance. Working Paper, Stanford University, 2015.

G. Ellison. Bounded rationality in industrial organization, in *Advances in Economics and Econometrics: Theory and Applications,* eds Blundell R., Newey W., and Persson T. Cambridge University Press, 2006.

W. Elmaghraby and S. Oh. Procurement auctions and eroding price contracts in the presence of learning by doing. Working Paper, University of Maryland College Park, 2013.

B. Elyakime, J. J. Laffont, P. Loisel, and Q. Vuong. First-price sealed-bid auctions with secret reservation prices. *Annals of Economics and Statistics*, 34:115–141, 1994.

A. Farahat and G. Perakis. A comparison of bertrand and cournot profits in oligopolies with differentiated products. *Operations Research*, 59:507–513, 2011.

A. Federgruen and M. Hu. Price competition in sequential multi-product oligopolies. Working paper, Columbia University, New York, NY, 2012.

A. Federgruen and L. Lu. Price competition based on relative prices. Working Paper, Columbia Business School, 2016a.

A. Federgruen and L. Lu. Medicare reform: Estimation of the impacts of premium support systems. Working Paper, Columbia Business School, 2016b.

R. Feldman, R. Coulam, and B. Dowd. Competitive bidding can help solve medicare's fiscal crisis. http://www.aei.org/outlook/health/healthcare-reform/competitive-bidding-can-help-solve-medicares-fiscal-crisis/. *American Enterprise Institute.*, 2012.

Q. Feng and X. Lu. Supply chain contracting under competition: Bilateral bargaining vs. stackelberg. *Production and Operations Management*, 22:661–675, 2013.

G. Fibich and N. Gavish. Numerical simulations of asymmetric first-price auctions. *Games and Economic Behavior*, 73:479–495, 2011.

P. Fuloria and S. Zenios. Outcomes-adjusted reimbursement in a health-care delivery system. *Management Science*, 47:735–751, 2001.

G. Gallego and R. Wang. Multi-product price optimization and competition under the nested logit model with product-differentiated price sensitivities. *Operations Research*, 62:450–461, 2014.

G. Gallego, T. Huh, W. Kang, and R. Phillips. Price competition with the attraction demand model: Existence of unique equilibrium and its stability. *Manufacturing & Service Operations Management*, 8:359–375, 2006.

J. K. Goeree and T. Offerman. Competitive bidding in auctions with private and common values. *The Economic Journal*, 113:598–613, 2003.

F. Gonul and A. Smith. A method to infer coupon availability from coupon redemption in the supermarket scanner panel data. *Journal of Retailing and Consumer Services*, 6:107–115, 1999.

A. Ha and S. Tong. Contracting and information sharing under supply chain competition. *Management Science*, 54:701–715, 2008.

A. Ha, S. Tong, and H. Zhang. Sharing demand information in competing supply chains with production diseconomies. *Management Science*, 57:566–581, 2011.

J. Hadar. A note on dominant diagonals in stability analysis. *Econometrica*, 33:442–444, 1965.

A. Hall. The value of medicare managed care plans and their prescription drug benefits. Finance and Economics Discussion Series, Federal Reserve Board of Governers, Washington, D.C., 2007.

P. Heidhues and B. Kőszegi. Competition and price variation when consumers are loss averse. *American Economic Review*, 98:1245–1268, 2008.

T. Hubbard and J. Paarsch. On the numerical solution of equilibria in auction models with asymmetries within the private-values paradigm. Working paper, Colby College, 2011.

A. Jaeger-Waldau. Pv status report 2007. JRC Technical Notes, http://www.pv-era.net/doc_upload/documents/221_0086PVStatusReport2007.pdf, 2007.

H. Jiang, Z. Pang, and S. Savin. Performance-based contracts for outpatient medical services. *About Manufacturing and Service Operations Management*, 14:654–669, 2012.

K. Johnson-Ferreira, A. Lee, and D. Simchi-Levi. Analytics for an online retailer: Demand forecasting and price optimization. To appear in Manufacturing and Service Operations Management, 2016.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

G-H Kang. Korean multiple award schedule. In *Presentation at 2013 9th Ninth Annual Conference of the INGP*, 2013.

G. Kannan and M. Popiuc. Reverse supply chain coordination by revenue sharing contract: A case for the personal computers industry. *European Journal of Oper. Res.*, 233:326–336, 2014.

H. Karle and M. Peitz. Competition under consumer loss aversion. *The RAND Journal of Economics*, 45:1–31, 2014.

G. King, X. Chao, and I. Duenyas. Who benefits when prescription drug manufacturers offer copay coupons? Working Paper, University of Michigan, 2014.

B. Köszegi and M. Rabin. A model of reference-dependent preferences. *The Quarterly Journal of Economics*, 4:1133–1165, 2006.

V. Krishna. *Auction Theory.* Academic Press, 2002.

H. Krishnan, R. Kapuscinski, and D. A. Butz. Coordinating contracts for decentralized supply chains with retailer promotional effort. *Management Sci.*, 50:48–63, 2004.

M. Lariviere and E. Porteus. Selling to the newsvendor: An analysis of price-only contracts. *Manufacturing Service Oper. Management*, 3:293–305, 2001.

D. Lee and S. Zenios. An evidence-based incentive system for medicare's end-stage renal disease program. *Management Science*, 58:1092–1105, 2012.

R. Levi, G. Perakis, and G. Romero. On the effectiveness of uniform subsidies in increasing market consumption. To appear in Management Science, 2016.

H. Li and W. Huh. Pricing multiple products with the multinomial logit and nested logit models: Concavity and implications. *Manufacturing Service Operations Management*, 13:549–564, 2011.

G. Liu. On nash equilibrium in prices in an oligopolistic market with demand characterized by a nested multinomial logit model and multiproduct firm as nest. Discussion Papers, Statistics Norway, 2006.

R. Lobel and G. Perakis. Consumer choice model for forecasting demand and designing incentives for solar technology. Working Paper, Wharton School, 2016.

M. Lu and C. Donaldson. Performance-based contracts and provider efficiency: The state of the art. *Disease Management and Health Outcomes*, 7:127–137, 2000.

L. Lu and Y. Wu. Harmony in competition: On preferences for contractual forms in supply chains. Working paper, Columbia Business School, 2013.

J. Lustig. The welfare effects of adverse selection in privatized medicare. Working Paper, Boston University, 2008.

A. Mas-Colell, M. Whinston, and J. Green. *Microeconomic Theory*. Oxford University Press, 1995.

T. Mazumdar and P. Papatla. An investigation of reference price segments. *Journal of Marketing Research*, 37:246–258, 2000.

T. Mazumdar, S. Raj, and I. Sinha. Reference price research: Review and propositions. *Journal of Marketing*, 69:84–102, 2005.

T. McGuire and R. Staelin. An industry equilibrium analysis of downstream vertical integration. *Marketing Sci.*, 2:161–191, 1983.

C. Mela, S. Gupta, and D. Lehmann. The long term impact of promotion and advertising on consumer brand choice. *Journal of Marketing Research*, 34:248–261, 1997.

P. Milgrom. *Putting auction theory to work*. Cambridge University Press, 2004.

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

M. Nagarajan and G. Sosic. Game-theoretic analysis of cooperation among supply chain agents: Review and extensions. *European Journal of Oper. Res.*, 187:719–745, 2008.

S. Netessine and F. Zhang. Positive vs. negative externalities in inventory management: Implications for supply chain design. *Manufacturing Service Oper. Management*, 7:58–73, 2005.

K. Nosal. Estimating switching costs for medicare advantage plans. Working paper, University of Mannheim, Germany, 2012.

B. Pasternack. Optimal pricing and returns policies for perishable commodities. *Marketing Sci.*, 4:166–176, 1985.

A. Rivlin and W. Daniel. Could improving choice and competition in medicare advantage be the future of medicare. Washington, D.C.: The Brookings Institution, http://www.brookings.edu/~/media/Research/Files/Papers/2015/06/04-medicare-2030-paper-series/060315RivlinDanielMedicareAdvantage.pdf?la=en, 2015.

D. Saban and G.Y. Weintraub. Procurement mechanisms for differentiated products. Working Paper, 2015.

J. Schummer and R. Vohra. Auctions for procuring options. *Operations Research*, 51:41–51, 2003.

K. So and C. Tang. Modeling the impact of an outcome-oriented reimbursement policy on clinic, patients, and pharmaceutical firms. *Management Science*, 46:875–892, 2000.

Z. Song, D. Cutler, and M. Chernew. Potential consequences of reforming medicare into a competitive bidding system. *The Journal of the American Medicine Association*, 308:459–460, 2012a.

Z. Song, M. Landrum, and M. Chernew. Competitive bidding in medicare: Who benefits from competition? *The American Journal of Managed Care*, 18:546–555, 2012b.

Z. Song, M. Landrum, and M. Chernew. Competitive bidding in medicare advantage: Effect of benchmark changes on plan bids. *Journal of Health Economics*, 32:1301–1312, 2013.

R. Spiegler. *Bounded Rationality and Industrial Organization.* Oxford University Press, New York, 2011.

J. Sun and L. Debo. Sustaining long-term supply chain partnerships using price–only contracts. *European Journal of Oper. Res.*, 223:557–565, 2014.

T. Taylor. Coordination under channel rebates with sales effort effect. *Management Sci.*, 48:992–1007, 2002.

J. Tirole. *The Theory of Industrial Organization.* The MIT Press, Cambridge, MA, 1988.

B. Tomlin. Capacity investments in supply chains: Sharing the gain rather than sharing the pain. *Manufacturing Service Oper. Management*, 5:317–333, 2003.

D. Topkis. *Supermodularity and Complementarity.* Princeton University Press, Princeton, NJ, 1998.

A. Tsay. Quantity-flexibility contract and supplier-customer incentives. *Management Sci.*, 45:1339–1358, 1999.

T. I. Tunca and S. A. Zenios. Supply auctions and relational contracts for procurement. *Manufacturing & Service Operations Management*, 8:43–67, 2006.

A. Tversky and D. Kahneman. Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106:1039–1061, 1991.

X. Vives. *Oligopoly pricing: Old ideas and new tools.* The MIT Press, 2001.

Wall Street Journal. This week. `http://online.wsj.com/news/articles/SB10001424052748703338004575230240755797572`, May 8, 2010.

R. Winer. A reference price model of brand choice. *Journal of Marketing Research*, 13:250–256, 1986.

L. Yang, de F. Vericourt, and P. Sun. Time-based competition with benchmark effects. *Manufacturing & Service Operations Management*, 16:119–132, 2014.

F. Zhang. Procurement mechanism design in a two-echelon inventory system with price-sensitive demand. *Manufacturing & Service Operations Management*, 12:608–626, 2010.

J. Zhou. Reference dependence and market competition. *Journal of Economics and Management Strategy*, 20:1073–1097, 2011.

# Appendix A

# Appendix for Chapter 1

**Proof of Lemma 1.** It is obvious that $\beta \leq \alpha$ from their definitions. Fix $i = 0, 1, \ldots, n$, since $g_i(p)$ is a continuous function, there are finite values $L_i < U_i$, such that $-\infty < L_i \equiv \min_{p^{min} \leq p \leq p^{max}} \{p_i - g_i(p)\} < \max_{p^{min} \leq p \leq p^{max}} \{p_i - g_i(p)\} \equiv U_i < \infty$. Let $\{\xi_1, \xi_2, \ldots, \xi_m\}$ be the points in $\mathbb{P}$ that belong to $[L_i, U_i]$ and that $\xi_0 = L_i$ and $\xi_{m+1} = U_i$. Then,

$$\max\{f'(p_i - g_i(\mathbf{p})) : \ p \in [p^{min}, p^{max}] \text{ and } p_i - g_i(\mathbf{p})) \notin \mathbb{P}\} \leq \max_{r=0,1,\ldots,m} \max_{\xi_r \leq u \leq \xi_{r+1}} \{f'(u)\}, \quad \text{(A-1)}$$

where, at any left (right) end point of an interval $[\xi_r, \xi_{r+1}]$, $f'(u)$ is replaced by a right- or left-hand derivative, thus establishing continuous extensions. It follows that the right hand side of (A-1) is bounded and hence $\alpha < \infty$. The proof that $\beta > 0$ is analogous, creating similar continuous extension on the partition $\{\xi_0, \xi_1, \ldots, \xi_{m+1}\}$. □

**Proof of Theorem 1.** We first prove the theorem when $f$ is piecewise linear; we then give the proof for a general function $f$ that satisfies Assumption 1, by approximating this function by a sequence of piecewise linear functions,

**Part 1:** Assume, the function $f(\cdot)$, is piecewise linear with $M + 1$ segments, characterized by

$M$ break points $(x_1, x_2, \ldots, x_M)$ and $M+1$ non-negative slopes $(\beta_1, \beta_2, \ldots, \beta_{M+1})$ as follows:

$$f(x) = \begin{cases} f(x_1) + \beta_1(x - x_1), & x \le x_1 \\ f(x_1) + \beta_2(x - x_1), & x \in [x_1, x_2] \\ \vdots & \vdots \\ f(x_{M-1}) + \beta_M(x - x_{M-1}), & x \in [x_{M-1}, x_M] \\ f(x_M) + \beta_{M+1}(x - x_M), & x \ge x_M \end{cases}$$

For any $i = 1, 2, \ldots, N$, and fixed prices $p_{i1} > p_{i2}$, we show that the difference in the logarithms of the profit function for firm $i$, $\log(\pi_i(p_{i1}, p_{-i})) - \log(\pi_i(p_{i2}, p_{-i}))$, is non-decreasing in any competitor's price $p_j, j \neq i$ for any $p_{-i}$. It suffices to show this property is satisfied in the following two cases: (i) $\Delta p_{i2} < \Delta p_{i1}$ are contained in the same line segment, i.e., $x_{m-1} \le \Delta p_{i2} < \Delta p_{i1} \le x_m$ for some $m = 1, 2, \ldots, M+1$; (ii) $\Delta p_{i2}, \Delta p_{i1}$ are separated by exactly one break point $x_m$, i.e., $\Delta p_{i2} < x_m < \Delta p_{i1}$ for some $m = 1, 2, \ldots, M$. (If $\Delta p_{i2}, \Delta p_{i1}$ are separated by more than one break point, the difference $\log(\pi_i(p_{i1}, p_{-i})) - \log(\pi_i(p_{i2}, p_{-i}))$ can be written as the sum of differences involving pairs of price levels that are separated by a single break point.) [1]

(i) $\underline{x_{m-1} \le \Delta p_{i2} < \Delta p_{i1} \le x_m}$. Note that

$$\log(\pi_i(p_{i1}, p_{-i})) - \log(\pi_i(p_{i2}, p_{-i})) = \int_{p_{i2}}^{p_{i1}} \frac{\partial \log(\pi_i)}{\partial p_i}(p_i, p_{-i}) dp_i,$$

since for all $p_i \in (p_{i1}, p_{i2})$, $\Delta p_i = p_i - C$ is in the interior of the same line segment, where the function $f(\cdot)$ is differentiable. Hence, it is sufficient to show that $\frac{\partial \log \pi_i(p_i, p_{-i})}{\partial p_i}$ is non-decreasing in $p_j, j \neq i$, for any $p_i \in (C + x_{m-1},\ C + x_m)$. For any $p_i \in (C + x_{m-1},\ C + x_m)$, we have

$$\frac{\partial \log(\pi_i)}{\partial p_i} = \frac{1}{p_i - c_i} + \frac{\partial d_i}{\partial p_i} / d_i = \frac{1}{p_i - c_i} - b_i f'(p_i - C)(1 - d_i).$$

Since $f'(p_i - C) = \beta_m$ for any $p_i \in (C + x_{m-1},\ C + x_m)$ and $d_i$ is non-decreasing in $p_j, j \neq i$, it follows that $\frac{\partial \log(\pi_i)}{\partial p_i}$ is non-decreasing in $p_j, j \neq i$.

---

[1] $x_0 = -\infty, x_{M+1} = +\infty$.

(ii) $\underline{C + x_{m-1} < p_{i2} < C + x_m < p_{i1} < C + x_{m+1}}$. One has, for $\delta > 0$ sufficiently small, that

$$\begin{aligned}\log \pi_i(p_{i1}, p_{-i}) - \log \pi_i(p_{i2}, p_{-i}) &= \overbrace{\log \pi_i(p_{i1}, p_{-i}) - \log \pi_i(C + x_m + \delta, p_{-i})} \\ &\quad + \underline{\log \pi_i(C + x_m + \delta, p_{-i}) - \log \pi_i(C + x_m - \delta, p_{-i})} \\ &\quad + \underbrace{\log \pi_i(C + x_m - \delta, p_{-i}) - \log \pi_i(p_{i2}, p_{-i})}.\end{aligned}$$

By case (i), it suffices to show that there exists $\delta > 0$ small enough such that

$$\log \pi_i(C + x_m + \delta, p_{-i}) - \log \pi_i(C + x_m - \delta, p_{-i}) \text{ is non-decreasing in } p_j, j \neq i. \tag{A-2}$$

We will show this by considering two cases

Case (ii.a): For any $\Delta p_j \notin \mathbb{P}$, let

$$\Delta_j(\delta) = \frac{\partial \left[\log \pi_i(C + x_m + \delta, p_{-i}) - \log \pi_i(C + x_m - \delta, p_{-i})\right]}{\partial p_j}.$$

In the following, we will show for any $j \neq i$,

$$\Delta_j(\delta) \geq 0 \quad \text{for any } \delta > 0. \tag{A-3}$$

By the definition of $\Delta_j(\delta)$, one has

$$\begin{aligned}\Delta_j(\delta) &= \frac{\partial \left[\log \pi_i(C + x_m + \delta, p_{-i}) - \log \pi_i(C + x_m - \delta, p_{-i})\right]}{\partial p_j} \\ &= \frac{\partial \log d_i(C + x_m + \delta, p_{-i})}{\partial p_j} - \frac{\partial \log d_i(C + x_m - \delta, p_{-i})}{\partial p_j} \\ &= \frac{b_j f'(p_j - C) \exp(a_j - b_j f(p_j - C))}{\exp(a_i - b_i f(x_m) - b_i \beta_{m+1}\delta) + \exp(a_j - b_j f(p_j - C)) + \sum_{k\neq i,j}^{N} \exp(a_k - b_k f(p_k - C))} \\ &\quad - \frac{b_j f'(p_j - C) \exp(a_j - b_j f(p_j - C))}{\exp(a_i - b_i f(x_m) + b_i \beta_m \delta) + \exp(a_j - b_j f(p_j - C)) + \sum_{k\neq i,j} \exp(a_k - b_k f(p_k - C))} \\ &= \frac{b_j f'(p_j - C) e^{a_j - b_j f(p_j - C)} \cdot \left[e^{a_i - b_i f(x_m) + b_i \beta_m \delta} - e^{a_i - b_i f(x_m) - b_i \beta_{m+1}\delta}\right]}{\left[e^{a_i - b_i f(x_m) - b_i \beta_{m+1}\delta} + \sum_{k\neq i} e^{a_k - b_k f(p_k - C)}\right] \cdot \left[e^{a_i - b_i f(x_m) + b_i \beta_m \delta} + \sum_{k\neq i} e^{a_k - b_k f(p_k - C)}\right]} \\ &= \frac{b_j f'(p_j - C) e^{a_j - b_j f(p_j - C)} e^{a_i - b_i f(x_m) + b_i \beta_m \delta} \cdot \left[1 - e^{-b_i(\beta_m + \beta_{m+1})\delta}\right]}{\left[e^{a_i - b_i f(x_m) - b_i \beta_{m+1}\delta} + \sum_{k\neq i} e^{a_k - b_k f(p_k - C)}\right] \cdot \left[e^{a_i - b_i f(x_m) + b_i \beta_m \delta} + \sum_{k\neq i} e^{a_k - b_k f(p_k - C)}\right]} \\ &\geq 0, \qquad \text{for any } \delta > 0.\end{aligned}$$

Case (ii.b). In this case, we show (A-2) holds when $\Delta p_j \in \mathbb{P}$. Fixing $p_{-ij}$, we simplify the notation by writing $d_i(p_i, p_j)$ and $\pi_i(p_i, p_j)$ instead of $d_i(p_i, p_{-i})$ and $\pi_i(p_i, p_{-i})$. For $\delta > 0$ small enough, we will show

$$\begin{aligned}&\log \pi_i(C + x_m + \delta, C + x_o + \delta) - \log \pi_i(C + x_m + \delta, C + x_o - \delta) \\ &\geq \log \pi_i(C + x_m - \delta, C + x_o + \delta) - \log \pi_i(C + x_m - \delta, C + x_o - \delta).\end{aligned} \tag{A-4}$$

Substituting the demand equation (1.10) into the profit function, we have

$$\begin{aligned}
& \log \pi_i(C + x_m + \delta, C + x_o + \delta) - \log \pi_i(C + x_m + \delta, C + x_o - \delta) \\
& - [\log \pi_i(C + x_m - \delta, C + x_o + \delta) - \log \pi_i(C + x_m - \delta, C + x_o - \delta)] \\
= \quad & \log d_i(C + x_m + \delta, C + x_o + \delta) + \log d_i(C + x_m - \delta, C + x_o - \delta) \\
& - \log d_i(C + x_m + \delta, C + x_o - \delta) - \log d_i(C + x_m - \delta, C + x_o + \delta) \\
= \quad & \log \left( \frac{d_i(C + x_m + \delta, C + x_o + \delta)}{d_i(C + x_m + \delta, C + x_o - \delta)} \right) + \log \left( \frac{d_i(C + x_m - \delta, C + x_o - \delta)}{d_i(C + x_m - \delta, C + x_o + \delta)} \right) = \log \left( \frac{A_1 A_2}{B_1 B_2} \right),
\end{aligned}$$

where

$$\begin{aligned}
A_1 &= \left( e^{a_i - b_i f(x_m) - b_i \beta_{m+1} \delta} + e^{a_j - b_j f(x_o) + b_j \beta_o \delta} + \sum_{k \neq i,j} e^{a_k - b_k f(p_k - C)} \right), \\
B_1 &= \left( e^{a_i - b_i f(x_m) - b_i \beta_{m+1} \delta} + e^{a_j - b_j f(x_o) - b_j \beta_{o+1} \delta} + \sum_{k \neq i,j} e^{a_k - b_k f(p_k - C)} \right), \\
A_2 &= \left( e^{a_i - b_i f(x_m) + b_i \beta_m \delta} + e^{a_j - b_j f(x_o) - b_j \beta_{o+1} \delta} + \sum_{k \neq i,j} e^{a_k - b_k f(p_k - C)} \right), \\
B_2 &= \left( e^{a_i - b_i f(x_m) + b_i \beta_m \delta} + e^{a_j - b_j f(x_o) + b_j \beta_o \delta} + \sum_{k \neq i,j} e^{a_k - b_k f(p_k - C)} \right).
\end{aligned}$$

Let $\Gamma_i(\delta) \equiv A_1 A_2 - B_1 B_2$. Hence (A-4) is equivalent to $\Gamma_i(\delta) \geq 0$. Since $\Gamma_i(0) = 0$, to show $\Gamma_i(\delta) \geq 0$ for $\delta > 0$ sufficiently small, it suffices to show $\lim_{\delta \searrow 0} \Gamma_i'(\delta) = 0$ and $\lim_{\delta \searrow 0} \Gamma_i''(\delta) > 0$. Indeed, taking derivatives w.r.t $\delta$, it can be shown that

$$\begin{aligned}
\Gamma_i'(\delta) \quad = \quad & e^{a_i - b_i f(x_m) + a_j - b_j f(x_o) - (\beta_{m+1} b_i + \beta_{o+1} b_j) \delta} \cdot \Big[ - \beta_m b_i e^{(\beta_m + \beta_{m+1}) b_i \delta} - \beta_o b_j e^{(\beta_o + \beta_{o+1}) b_j \delta} \\
& + (\beta_m b_i + \beta_o b_j) e^{((\beta_m + \beta_{m+1}) b_i + (\beta_o + \beta_{o+1}) b_j) \delta} + \beta_{o+1} b_j (-1 + e^{(\beta_m + \beta_{m+1}) b_i \delta}) \\
& + \beta_{m+1} b_i (-1 + e^{(\beta_o + \beta_{o+1}) b_j \delta}) \Big],
\end{aligned}$$

so that

$$\lim_{\delta \searrow 0} \Gamma_i'(\delta) = 0 \quad \text{and} \quad \lim_{\delta \searrow 0} \Gamma_i''(\delta) = 2 b_i b_j (\beta_o + \beta_{o+1})(\beta_m + \beta_{m+1}) e^{a_i - b_i f(x_m) + a_j - b_j f(x_o)} > 0.$$

Hence, we have shown that for $\delta > 0$ small enough, (A-4) holds, so that, $\log(\pi_i)$ is supermodular.

**Part 2:** Assume now $f$ is a general continuous and increasing function. It is well-known that there exists a sequence of increasing piece-wise linear functions $\{f^{(k)}(\cdot)\}$ such that $\lim_{k \to \infty} f^{(k)}(x) = f(x)$ for any $x$. For any $i$, let $\pi_i^{(k)}$ denote firm $i$'s profit function associated with the function $f^{(k)}$. For any given pair of price vectors, $\mathbf{p}, \mathbf{p}'$, such that $\mathbf{p} \geq \mathbf{p}'$, we have by part 1, that:

$$\log \pi_i^{(k)}(p_i, p_{-i}) - \log \pi_i^{(k)}(p_i', p_{-i}) \geq \log \pi_i^{(k)}(p_i, p_{-i}') - \log \pi_i^{(k)}(p_i', p_{-i}'). \tag{A-5}$$

By a simple continuity argument, we have, for any price vector $\hat{p}$, $\lim_{k\longrightarrow\infty} \log \pi_i^{(k)}(\hat{p}) = \log \pi_i(\hat{p})$. Hence, taking limits in (A-5) yields

$$\log \pi_i(p_i, p_{-i}) - \log \pi_i(p_i', p_{-i}) \geq \log \pi_i(p_i, p_{-i}') - \log \pi_i(p_i', p_{-i}'),$$

thus, showing that the price game is log-supermodular for a general convex increasing function $f$.

□

**Proof of Proposition 1.** Analogous to the proof of Theorem 1, we first establish the Proposition when $f$ is piecewise linear; we then extend the proof for a general convex and increasing function $f$, by approximating this function by a sequence of piecewise linear functions,

**Part 1:** Assume, the function $f(\cdot)$ is piecewise linear, as specified in Part 1 of the proof of Theorem 1, with $0 \leq \beta_1 \leq \cdots \leq \beta_{M+1}$ in view of convexity. For any firm $i = 1, 2, \ldots, N$, we show that the logarithms of the profit function, $\log(\pi_i)$, is quasi-concave. In what follows, we show this property by considering two cases: (1) $\Delta p_i \notin \mathbb{P}$; (2) $\Delta p_i = x_m$ for some $m = 1, 2, \ldots, M$.

(i) $\underline{\Delta p_i \notin \mathbb{P}}$, say, $\Delta p_i \in (x_{m-1}, x_m)$ for some $m = 1, 2, \ldots, M+1$. Note that, the function $f(\Delta p_i)$ is differentiable and, by (1.11),

$$\frac{\partial \log(\pi_i)}{\partial p_i} = \frac{1}{p_i - c_i} + \frac{\partial d_i}{\partial p_i}/d_i = \frac{1}{p_i - c_i} - b_i f'\left(\Delta p_i\right)(1 - d_i).$$

Since $f'\left(\Delta p_i\right) = \beta_m$ when $\Delta p_i \in (x_{m-1}, x_m)$ and $d_i$ is non-increasing in $p_i$, by (1.11), it follows that $\frac{\partial \log(\pi_i)}{\partial p_i}$ is non-increasing in $p_i$.

(ii) $\underline{\Delta p_i = x_m \text{ for some } m = 1, 2, \ldots, M.}$ We show that $\frac{\partial_+ \log \pi_i(C+x_m, p_{-i})}{\partial p_i} \leq \frac{\partial_- \log \pi_i(C+x_m, p_{-i})}{\partial p_i}$, which is equivalent to $\frac{\partial_+ \log d_i(C+x_m, p_{-i})}{\partial p_i} \leq \frac{\partial_- \log d_i(C+x_m, p_{-i})}{\partial p_i}$, or, by (1.11),

$$\begin{aligned}
\frac{\partial_+ \log d_i(C+x_m, p_{-i})}{\partial p_i} &\equiv \lim_{p_i \searrow C+x_m} \frac{\partial \log d_i(p_i, p_{-i})}{\partial p_i} = -b_i f_+'(x_m)\left(1 - d_i(C+x_m, p_{-i})\right) \\
&= -b_i \beta_{m+1}\left(1 - d_i(C+x_m, p_{-i})\right) \\
&\leq -b_i \beta_m \left(1 - d_i(C+x_m, p_{-i})\right) = -b_i f_-'(x_m)\left(1 - d_i(C+x_m, p_{-i})\right) \\
&= \lim_{p_i \nearrow C+x_m} \frac{\partial \log d_i(p_i, p_{-i})}{\partial p_i} \equiv \frac{\partial_- \log d_i(C+x_m, p_{-i})}{\partial p_i}.
\end{aligned}$$

Here the inequality follows from the fact that $0 \leq \beta_m \leq \beta_{m+1}$ and $1 - d_i(C+x_m, p_{-i}) \geq 0$.

We have thus shown that $\log \pi_i(p_i, p_{-i})$ is quasi-concave in $p_i \geq c_i$ for any $p_{-i}$ when the response function $f$ is increasing and convex piecewise linear.

**Part 2:** Assume now that $f$ is a general increasing and convex function. It is well-known that there exists a sequence of increasing and convex piece-wise linear functions $\{f^{(k)}(\cdot)\}$ such that $\lim_{k\to\infty} f^{(k)}(x) = f(x)$ for any $x$. We prove the result by contradiction. Assume that there exists some firm $i = 1, 2, \dots, N$ and price vector $p^0_{-i}$ such that $\log(\pi_i(\cdot, p^0_{-i}))$ is not quasi-concave, that is,

$$\log \pi_i \left(\lambda_0 p_i^1 + (1-\lambda_0) p_i^2, p^0_{-i}\right) < \min\left\{\log \pi_i \left(p_i^1, p^0_{-i}\right), \log \pi_i \left(p_i^2, p^0_{-i}\right)\right\}, \tag{A-6}$$

for some $p_i^1, p_i^2$ and $\lambda_0 \in (0,1)$. Let $\pi_i^{(k)}$ denote firm $i$'s profit function associated with the function $f^{(k)}$. By a simple continuity argument, we have, for any price vector $\hat{p}$, $\lim_{k\longrightarrow\infty} \log \pi_i^{(k)}(\hat{p}) = \log \pi_i(\hat{p})$. Hence, by (A-6), there exists $k_0 \geq 1$ such that

$$\log \pi_i^{(k)} \left(\lambda_0 p_i^1 + (1-\lambda_0) p_i^2, p^0_{-i}\right) < \min\left\{\log \pi_i^{(k)} \left(p_i^1, p^0_{-i}\right),\ \log \pi_i^{(k)} \left(p_i^2, p^0_{-i}\right)\right\} \quad \text{for any } k \geq k_0.$$

This contradicts the quasi-concavity of $\pi_i^{(k)}$ in part 1. □

**Proof of Theorem 2.** In Theorem 1, we show that each of the profit functions $\{\pi_i : i = 1, 2, \dots, N\}$ is log-supermodular. By Theorem 4.3.2 in Topkis [1998], it therefore suffices to prove that for each $i = 1, 2, \dots, N$, $\tilde{\pi}_i = \log(\pi_i)$ is supermodular in the pair $(p_i, C)$, or the difference function $\tilde{\pi}_i(p_i + \Delta, p_{-i}) - \tilde{\pi}_i(p_i, p_{-i})$, for any $\Delta > 0$, is an increasing function of $C$. Similar to the proof of Theorem 1, by Assumption 1, it follows that

$$\tilde{\pi}_i(p_i + \Delta, p_{-i}) - \tilde{\pi}_i(p_i, p_{-i}) \quad = \quad \int_{p_i}^{p_i+\Delta} \frac{\partial \tilde{\pi}_i}{\partial u}(u, p_{-i}) du, \tag{A-7}$$

where in the finitely many points in which the partial derivative in the integral fails to exists, it is replaced by either the right-hand or the left-hand derivative. When it exists, by (1.11) and (1.12), one has

$$\frac{\partial \tilde{\pi}_i}{\partial p_i} = \frac{1}{p_i - c_i} - b_i f'(p_i - C)(1 - d_i)$$

Since $f$ is convex, by (A-7), it suffices to show that

For any price vector $p$, $d_i(p)$ is almost everywhere non-decreasing in $C$, for all $i = 1, 2, \dots, N$ (A-8)

This, however, follows immediately from the dominant-diagonal condition (D), since a marginal increase of $C$ by a quantity $\Delta$ is equivalent to a simultaneous decrease of all prices by the same amount $\Delta$. For $\Delta$ sufficiently small, the sign of the change in the sales volume $d_i$ is given by $\sum_{j=0}^{N} \frac{\partial d_i}{\partial p_j} \leq 0$, by (D). □

**Proof of Proposition 2.** In view of the proof of Theorem 2, it suffices to show that (1.17) holds, since as argued, (1.17) is equivalent to (D). To verify (1.17), note that, by (1.11) and (1.12), where the partial derivative exists

$$\begin{aligned} |\frac{\partial d_i}{\partial p_i}| - \sum_{j \neq i} \frac{\partial d_i(p)}{\partial p_j} &= bf'(p_i - C)d_i(1-d_i) - \sum_{j=1, j\neq i}^{N} bf'(p_j - C)d_i d_j \\ &= bd_i \left( f'(p_i - C) - \sum_{j=1}^{N} f'(p_j - C) d_j \right) \\ &\overset{(Eq)}{\geq} bd_i(\beta - \alpha \sum_{j=1}^{N} d_j) \\ &= bd_i \alpha \left( \frac{\beta}{\alpha} - 1 + d_0 \right) \geq 0, \end{aligned}$$

where the inequality (Eq) follows from the fact that $f'(x) \in [\beta, \alpha]$ for any $x$, and the last inequality follows by the assumption that $d_0 \geq 1 - \frac{\beta}{\alpha}$. □

**Proof of Theorem 3.** (a) The model may be viewed as a special case of the general attraction model addressed in Bernstein and Federgruen [2004], with the logarithm of firm $i$'s attraction value given by $\widetilde{a}_i(p_i) = a_i - b_i f((1-\delta_i)p_i)$, which is *decreasing* in $p_i$ since $f(\cdot)$ is increasing. Thus, the monotonicity condition (7) in Bernstein and Federgruen [2004] is satisfied, and the result follows from Theorem 2(a), there.

(b) The uniqueness result follows from Gallego *et al.* [2006], verifying that all conditions, there, are satisfied: conditions (A1) and (A3) are immediate; for (A2), note that the elasticity function $\eta_i(p_i) = -p_i \widetilde{a}'_i(p_i) = p_i b_i (1-\delta_i) f'((1-\delta_i)p_i)$ is a non-decreasing function since the function $f(\cdot)$ is convex. Conditions (B)-(C) are easily verified, as well. □

**Proof of Proposition 3.** It suffices to show that each of profit function $\log(\pi_i)$ is supermodular in the pair $(p_i, \delta)$. Since the response function $f$ is linear on the positive half line, and net prices are always non-negative, it is easily verified that the profit function, $\pi$, is twice order differentiable. Without loss of generality, set $\alpha = 1$ in the proof. Since

$$d_i = \frac{\exp\left(a_i - b_i\left((1-\delta)p_i\right)\right)}{\sum_{k=0}^{n} \exp\left(a_k - b_k\left((1-\delta)p_k\right)\right)}.$$

One has $\frac{\partial \log(\pi_i)}{\partial p_i} = \frac{1}{p_i - c_i} + \frac{\partial \log(d_i)}{\partial p_i} = \frac{1}{p_i - c_i} - b_i(1-\delta)(1-d_i)$ where the last equation follows by taking derivative of $d_i$ with respect to $p_i$ (which is similar to (1.11)). Taking derivative with respect to $\delta$,

$$\begin{aligned}
\frac{\partial^2 \log(\pi_i)}{\partial p_i \partial \delta} &= b_i(1-d_i) + b_i(1-\delta)\frac{\partial d_i}{\partial \delta} \\
&= b_i(1-d_i) + b_i(1-\delta) \cdot \left[ b_i p_i d_i (1-d_i) - \sum_{j \neq i} b_j p_j d_i d_j \right] \\
&= b_i(1-d_i) + b_i(1-\delta) d_i \cdot \left[ b_i p_i (1-d_i) - \sum_{j \neq i} b_j p_j d_j \right] \\
&\overset{(a)}{\geq} b_i(1-d_i) + b_i(1-\delta) d_i \cdot \left[ \underline{b}\underline{p}(1-d_i) - \overline{b}\overline{p}(1-d_i) \right] \\
&= b_i(1-d_i)\left[1 - (1-\delta) d_i (\overline{b}\overline{p} - \underline{b}\underline{p})\right] \\
&\overset{(b)}{\geq} b_i(1-d_i)\left[1 - (1-\delta)(\overline{b}\overline{p} - \underline{b}\underline{p})\right] \\
&\geq 0 \text{ if } \delta \geq 1 - \tfrac{1}{\overline{b}\overline{p} - \underline{b}\underline{p}},
\end{aligned}$$

where inequality (a) follows from the definition that $\underline{b} = \min\{b_i, i = 0, 1, 2, \ldots, n\}, \overline{b} = \max\{b_i, i = 0, 1, 2, \ldots, n\}, \underline{p} = \min\{p_i^{min}, i = 0, 1, 2, \ldots, n\}, \overline{p} = \max\{p_i^{max}, i = 0, 1, 2, \ldots, n\}$, and inequality (b) follows from the fact that $d_i \leq 1$.

**Proof of Proposition 4.** (a) immediately follows from (1.19) and (1.21) and the paragraph proceeding Proposition 4.

(b) We distinguish between two cases: (i) $\underline{p_i \leq p_{(1)}^{-i}}$: the monotone property is immediate from (1.18). (ii) $\underline{p_i \geq p_{(1)}^{-i}}$: By (1.20), the monotonicity of $d_i$ with respect to $p_j$ is immediate when $p_j > p_{(1)}^{-i}$. When $p_j = p_{(1)}^{-i}$,

$$d_i(p_i, p_{-i}) = \frac{\exp\left(a_i - bf\left(p_i - p_j\right)\right)}{\exp\left(a_i - bf\left(p_i - p_j\right)\right) + \sum_{k \neq i} \exp\left(a_k - bf\left(p_k - p_j\right)\right)}, \tag{A-9}$$

and the expression remains valid when $p_j$ is decreased downward from $p_{(1)}^{-i}$. Thus,

$$\begin{aligned}\frac{\partial_- d_i}{\partial p_j} &= bf'(p_i - p_j)d_i - bd_i \sum_{k\neq j} f'_+(p_k - p_j)d_k \\ &= b\alpha d_i - b\alpha d_i \sum_{k\neq j} d_k = b\alpha d_i \left(1 - \sum_{k\neq j} d_k\right) \geq 0. \quad \text{whenever } p_j = p_{(1)}^{-i}. \end{aligned} \tag{A-10}$$

Similarly, when $p_j$ is the *unique* lowest price, i.e., $p_j < p_k$ for all $k \neq j$, (A-9) continues to apply even when $p_j$ is increased *upward* from $p_{(1)}^{-i}$, so that $\frac{\partial_+ d_i}{\partial p_j} = \frac{\partial_- d_i}{\partial p_j} = b\alpha d_i \left(1 - \sum_{k\neq j} d_k\right) \geq 0$. The remaining case has:

$$p_j = p_{(1)}^{-i} = p_l \quad \text{for some } l \neq j, i.$$

In view of (A-10), it suffices to show that $\frac{\partial_+ d_i}{\partial p_j} \geq 0$, which is immediate from the representation of $d_i$ in (1.20). □

**Proof of Theorem 4.** It suffices to show that the function $\pi_i(p_i, p_{-i})$ is quasi-concave in its own price for each firm $i = 1, 2, \ldots, N$. By (2.6), for firm $i = 1, 2, \ldots, N$, one has

$$\frac{\partial \pi_i}{\partial p_i} = d_i + (p_i - c_i)\frac{\partial d_i}{\partial p_i} = \begin{cases} d_i \left(1 - (p_i - c_i)\sum_{k\neq i} d_k b_k f'(p_k - p_i)\right), & p_i < p_{(1)}^{-i} \\ d_i \left(1 - (p_i - c_i) b_i f'\left(p_i - p_{(1)}^{-i}\right)(1 - d_i)\right), & p_i > p_{(1)}^{-i} \end{cases} \tag{A-11}$$

Also,

$$\frac{\partial_+ \pi_i(p_i, p_{-i})}{\partial p_i} = d_i + (p_i - c_i)\frac{\partial_+ d_i(p_i, p_{-i})}{\partial p_i}, \tag{A-12}$$

$$\frac{\partial_- \pi_i(p_i, p_{-i})}{\partial p_i} = d_i + (p_i - c_i)\frac{\partial_- d_i(p_i, p_{-i})}{\partial p_i}. \tag{A-13}$$

Thus, one has

$$\frac{\partial \pi_i}{\partial p_i} \geq 0 \quad \Leftrightarrow \quad \begin{cases} (p_i - c_i)\sum_{k\neq i} d_k b_k f'(p_k - p_i) \leq 1, & p_i < p_{(1)}^{-i} \\ (p_i - c_i) b_i f'\left(p_i - p_{(1)}^{-i}\right)(1 - d_i) \leq 1, & p_i > p_{(1)}^{-i} \end{cases} \tag{A-14}$$

Next, we will show quasi-concavity of $\pi_i(p_i, p_{-i})$ by showing that $\frac{\partial \pi_i}{\partial p_i} \geq 0$ if and only if $p_i \leq \hat{p}_i(p_{-i})$ for some threshold value $\hat{p}_i(p_{-i})$. In what follows, we show this property by considering three cases: (1) $p_i > p_{(1)}^{-i}$; (2) $p_i < p_{(1)}^{-i}$; and (3) $p_i = p_{(1)}^{-i}$.

(1) If $\underline{p_i > p_{(1)}^{-i}}$, $f'\left(p_i - p_{(1)}^{-i}\right)$ increases in $p_i$ since $f$ is convex on $[0, +\infty)$, $1 - d_i$ is increasing in $p_i$ since $\frac{\partial d_i}{\partial p_i} \leq 0$, therefore, $(p_i - c_i)b_i f'\left(p_i - p_{(1)}^{-i}\right)(1 - d_i)$ is increasing in $p_i$. Thus, once the function reaches a value $\geq 1$, i.e., once $\frac{\partial \pi_i(p_i, p_{-i})}{\partial p_i}$ is decreasing in a certain point, the same applies to any larger price value. This implies that the function $\pi_i(p_i, p_{-i})$ is quasi-concave on the interval $(p_{(1)}^{-i}, \infty)$.

(2) If $\underline{p_i < p_{(1)}^{-i}}$, the cross price derivatives of the demand functions are given by

$$\frac{\partial d_i}{\partial p_j} = b_j f'(p_j - p_i)\, d_i d_j, \text{ for any } j \neq i. \tag{A-15}$$

Let

$$H_i(p_i) = (p_i - c_i)\sum_{k\neq i} d_k b_k f'(p_k - p_i) - 1 = (p_i - c_i)\sum_{k\neq i} \frac{\partial d_i}{\partial p_k}/d_i - 1 = (p_i - c_i)\sum_{k\neq i} \frac{\partial \ln(d_i)}{\partial p_k} - 1.$$

Then, $\frac{\partial \pi_i}{\partial p_i} = -d_i H_i(p_i)$ by (A-11). Obviously, $H_i(c_i) = -1 < 0$ and we show quasi-concavity of $\pi_i$ by considering two cases: (2a) $H_i(p_{(1)}^{-i}-) \leq 0$; (2b) $H_i(p_{(1)}^{-i}-) > 0$.

Case (2a) $H_i(p_{(1)}^{-i}-) \leq 0$: then $H_i(p_i) \leq 0$ for all $p_i \in [c_i, p_{(1)}^{-i})$ since $H_i(c_i) < 0$ and $H_i(p_i)$ is quasi-convex by condition (M). Thus, $\frac{\partial \pi_i}{\partial p_i} = -d_i H_i(p_i) \geq 0$ for all $p_i \in [c_i, p_{(1)}^{-i})$.

Case (2b) $H_i(p_{(1)}^{-i}-) > 0$: by the quasi-convexity of $H_i(p_i)$ on $[c_i, p_{(1)}^{-i})$ and $H_i(c_i) < 0$, there exists $\hat{p}_i(p_{-i}) \in (c_i, p_{(1)}^{-i})$ such that $H_i(p_i) \leq 0$, thus, $\frac{\partial \pi_i}{\partial p_i} = -d_i H_i(p_i) \geq 0$, if and only if $p_i \leq \hat{p}_i(p_{-i})$.

Therefore, we have shown $\pi_i$ is quasi-concave in $p_i$ on the interval $[c_i, p_{(1)}^{-i})$ as well.

Thus, to complete the proof that the function $\pi_i(\cdot, p_{-i})$ is quasi-concave on the complete interval $[c_i, \infty)$, it suffices to show that $\frac{\partial_+ \pi_i}{\partial p_i}(p_{(1)}^{-i}, p_{-i}) \leq \frac{\partial_- \pi_i}{\partial p_i}(p_{(1)}^{-i}, p_{-i})$. By (A-12) and (A-13), it suffices to show $\frac{\partial_+ d_i(p_{(1)}^{-i}, p_{-i})}{\partial p_i} \leq \frac{\partial_- d_i(p_{(1)}^{-i}, p_{-i})}{\partial p_i}$, or

$$\begin{aligned}
\lim_{p_i \searrow p_{(1)}^{-i}} \frac{\partial d_i(p_i, p_{-i})}{\partial p_i} \equiv \frac{\partial_+ d_i(p_{(1)}^{-i}, p_{-i})}{\partial p_i} \quad &\leq \quad \frac{\partial_- d_i(p_{(1)}^{-i}, p_{-i})}{\partial p_i} = \lim_{p_i \nearrow p_{(1)}^{-i}} \frac{\partial d_i(p_i, p_{-i})}{\partial p_i} \\
&= \quad \lim_{p_i \nearrow p_{(1)}^{-i}} \left\{ -d_i \sum_{k\neq i} d_k b_k f'(p_k - p_i) \right\} = \lim_{p_i \nearrow p_{(1)}^{-i}} -\sum_{k\neq i} \frac{\partial d_i}{\partial p_k},
\end{aligned}$$

which is equivalent to (D). □

**Proof of Lemma 2.** Condition (D): For any $i = 1, 2, \ldots, N$, and any given $p_{-i}$, we have

$$\begin{aligned}
& \lim_{p_i \nearrow p_{(1)}^{-i}} \sum_{j \neq i} \frac{\partial d_i}{\partial p_j}(p_i, p_{-i}) \leq \lim_{p_i \searrow p_{(1)}^{-i}} |\frac{\partial d_i}{\partial p_i}(p_i, p_{-i})| \\
\Longleftrightarrow \quad & \lim_{p_i \nearrow p_{(1)}^{-i}} \left\{ \left( \sum_{j \neq i} d_i d_j b_j f'(p_j - p_i) \right) \right\} \leq \lim_{p_i \searrow p_{(1)}^{-i}} \left\{ b_i f' \left( p_i - p_{(1)}^{-i} \right) (1 - d_i) \, d_i \right\} \\
\overset{(a1)}{\Longleftrightarrow} \quad & d_i(p_{(1)}^{-i}, p_{-i}) \cdot \left( \sum_{j \neq i} d_j(p_{(1)}^{-i}, p_{-i}) b\alpha \right) \leq d_i(p_{(1)}^{-i}, p_{-i}) b\alpha \left( 1 - d_i(p_{(1)}^{-i}, p_{-i}) \right) \\
\Longleftrightarrow \quad & \sum_{j \neq i} d_j(p_{(1)}^{-i}, p_{-i}) \leq 1 - d_i(p_{(1)}^{-i}, p_{-i}),
\end{aligned}$$

which is trivially true. (a1) holds because $b_j = b$, $f'(p_j - p_i) = \alpha$ for any $p_i < p_{(1)}^{-i} = \min\{p_k, k \neq i\}$ and for any $j \neq i$.

Condition (M): For any $i = 1, 2, \ldots, N$, any given $p_{-i}$, and any $p_i \in [c_i, p_{(1)}^{-i})$, let

$$H_i(p_i) = (p_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial p_j} = (p_i - c_i) \sum_{j \neq i} d_j b_j f'(p_j - p_i) = (p_i - c_i) \sum_{j \neq i} d_j b\alpha.$$

Taking derivatives with respect to $p_i$, we get

$$\begin{aligned}
H_i'(p_i) &= \left[ \sum_{j \neq i} d_j + (p_i - c_i) \sum_{j \neq i} \frac{\partial d_j}{\partial p_i} \right] b\alpha \\
&\overset{(a2)}{=} \left[ \sum_{j \neq i} d_j + (p_i - c_i) \sum_{j \neq i} d_j \left( b\alpha - \sum_{k \neq i} d_k b\alpha \right) \right] b\alpha \\
&= \left[ 1 + b\alpha(p_i - c_i) \left( 1 - \sum_{j \neq i} d_j \right) \right] \cdot \left( \sum_{j \neq i} d_j \right) b\alpha \\
&\geq 0,
\end{aligned}$$

where (a2) holds because $d_j = \frac{\exp(a_j - bf(p_j - p_i))}{\exp(a_i) + \sum_{k \neq i} \exp(a_k - bf(p_k - p_i))} = \frac{\exp(a_j - b\alpha(p_j - p_i))}{\exp(a_i) + \sum_{k \neq i} \exp(a_k - b\alpha(p_k - p_i))}$ for any $p_i < p_{(1)}^{-i}$ and any $j \neq i$ by (1.18). Taking derivative w.r.t $p_i$ yields $\frac{\partial d_j}{\partial p_i} = d_j \left( b\alpha - \sum_{k \neq i} d_k b\alpha \right)$ for any $j \neq i$ and $p_i < p_{(1)}^{-i}$. Thus, $(p_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial p_j}$ is non-decreasing, which is quasi-convex, in $p_i$ on the interval $[c_i, p_{(1)}^{-i})$. □

**Proof of Theorem 5.** As in the proof of Theorem 1, for $p_{i1} > p_{i2}$, we show that the difference in the logarithms of firm $i$'s profit function, under $p_{i1}$ versus $p_{i2}$, is non-decreasing in $p_j, j \neq i$, for any $p_{-i}$. To this end, we distinguish among the following three cases: (i) $p_{i2} < p_{i1} \le p^{-i}_{(1)}$; (ii) $p^{-i}_{(1)} \le p_{i2} < p_{i1}$; (iii) $p_{i2} < p^{-i}_{(1)} < p_{i1}$.

(i) $\underline{p_{i2} < p_{i1} \le p^{-i}_{(1)}}$ or (ii) $\underline{p^{-i}_{(1)} \le p_{i2} < p_{i1}}$: Note that

$$\log \pi_i(p_{i1}, p_{-i}) - \log \pi_i(p_{i2}, p_{-i}) = \int_{p_{i2}}^{p_{i1}} \frac{\partial \log \pi_i(p_i, p_{-i})}{\partial p_i} dp_i,$$

since $\log \pi_i(p_i, p_{-i})$ is differentiable everywhere on the interval $(p_{i2}, p_{i1})$. Hence, it is sufficient to show that $\frac{\partial \log \pi_i(p_i, p_{-i})}{\partial p_i}$ is non-decreasing in $p_j$. By (1.19) and (1.21), we have

$$\frac{\partial \log \pi_i}{\partial p_i} = \begin{cases} \frac{1}{p_i - c_i} + \frac{\partial d_i}{\partial p_i}/d_i = \frac{1}{p_i - c_i} - b\sum_{k\neq i} f'(p_k - p_i)\, d_k = \frac{1}{p_i - c_i} - b\alpha(1-d_i), & \text{if (i) } p_i \le p^{-i}_{(1)} \\ \frac{1}{p_i - c_i} - bf'\left(p_i - p^{-i}_{(1)}\right)(1-d_i) = \frac{1}{p_i - c_i} - b\alpha(1-d_i), & \text{if (ii) } p_i > p^{-i}_{(1)} \end{cases}$$

which is non-decreasing in $p_j$ for any $j \neq i$, by Lemma 2 (b).

(iii) $\underline{p_{i2} < p^{-i}_{(1)} < p_{i1}}$: Similar to the proof of Theorem 1, fix $\delta \le \min\left\{p_{i1} - p^{-i}_{(1)}, p^{-i}_{(1)} - p_{i2}\right\}$,

$$\begin{aligned} \log \pi_i(p_{i1}, p_{-i}) - \log \pi_i(p_{i2}, p_{-i}) \quad = \quad & \overbrace{\log \pi_i(p_{i1}, p_{-i}) - \log \pi_i(p^{-i}_{(1)} + \delta, p_{-i})} \\ & + \underline{\log \pi_i(p^{-i}_{(1)} + \delta, p_{-i}) - \log \pi_i(p^{-i}_{(1)} - \delta, p_{-i})} \\ & + \underbrace{\log \pi_i(p^{-i}_{(1)} - \delta, p_{-i}) - \log \pi_i(p_{i2}, p_{-i})}. \end{aligned}$$

By cases (i) and (ii), both the first term and the third term are non-decreasing in $p_j$. It thus suffices to show that the second term, $\log \pi_i(p^{-i}_{(1)} + \delta, p_{-i}) - \pi_i(p^{-i}_{(1)} - \delta, p_{-i})$, is non-decreasing in $p_j, j \neq i$ as well. We show, in fact, that

$$\Delta_j(\delta) = \frac{\partial \left[\log \pi_i(p^{-i}_{(1)} + \delta, p_{-i}) - \log \pi_i(p^{-i}_{(1)} - \delta, p_{-i})\right]}{\partial p_j} \ge 0, \text{ for all } p_j,\ j \neq i. \tag{A-16}$$

Unless $p_j$ is the unique lowest price among firm $i$'s alternatives, so that the increase of $p_j$ is accompanied by an increase of $p^{-i}_{(1)}$, the proof of (A-16) is identical to the proof of case (ii.a) in Theorem 1. The remaining case for (A-16) has $p_j = p^{-i}_{(1)} < p_k$ for all $k \neq i, j$.

$$\begin{aligned} \Delta_j(\delta) \quad &= \quad \left.\frac{\partial \left[\log \pi_i(p_{i1}, p_{-i}) - \log \pi_i(p_{i2}, p_{-i})\right]}{\partial p_j}\right|_{p_{i1} = p^{-i}_{(1)} + \delta,\ p_{i2} = p^{-i}_{(1)} - \delta} \\ &= \quad \left.\frac{\partial \left[\log d_i(p_{i1}, p_{-i}) - \log d_i(p_{i2}, p_{-i})\right]}{\partial p_j}\right|_{p_{i1} = p^{-i}_{(1)} + \delta,\ p_{i2} = p^{-i}_{(1)} - \delta} \\ &= \quad b\alpha d_j\left(p^{-i}_{(1)} + \delta, p_{-i}\right) - b\alpha d_j\left(p^{-i}_{(1)} - \delta, p_{-i}\right) \ge 0. \quad \text{by Lemma 2 (b).} \end{aligned}$$

( The last equality follows from the fact that

$$\frac{\partial d_i(p_{i1}, p_{-i})}{\partial p_j} = b\alpha d_i \left(1 - \sum_{k \neq j} d_k\right) = b\alpha d_i d_j, \quad \text{for any } p_{i1} > p^{-i}_{(1)} = p_j \text{ by (A-10) in Appendix,}$$

$$\frac{\partial d_i(p_{i2}, p_{-i})}{\partial p_j} = bf'(p_j - p_{i2}) d_i d_j = b\alpha d_i d_j, \quad \text{for any } p_{i2} < p^{-i}_{(1)} = p_j \text{ by (1.18). )}$$

Thus, we have shown that the price competition game is log-supermodular.

□

**Proof of Theorem 6.** The proof is similar to that of Theorem 4; we show that each product $i$'s profit function is quasi-concave in its own price, i.e., the function has no local minimum. Note first from (1.24) that, if $p_i \neq p^{-i}_{(n-1)}, p^{-i}_{(n)}$, each firm $i$'s sales volume is differentiable in its own price, where

$$0 \geq \frac{\partial d_i}{\partial p_i} = \begin{cases} -b_i f'\left(p_i - p^{-i}_{(n-1)}\right) d_i(1 - d_i), & p_i < p^{-i}_{(n-1)} \\ -d_i \sum_{k \neq i} d_k b_k f'\left(p_k - p_i\right), & p_i \in (p^{-i}_{(n-1)}, p^{-i}_{(n)}) \\ -b_i f'\left(p_i - p^{-i}_{(n)}\right) d_i(1 - d_i), & p_i > p^{-i}_{(n)} \end{cases}, \tag{A-17}$$

Similar to (A-11), we therefore have

$$\frac{\partial \pi_i}{\partial p_i} = d_i + (p_i - c_i)\frac{\partial d_i}{\partial p_i} = \begin{cases} d_i \left(1 - (p_i - c_i) b_i f'\left(p_i - p^{-i}_{(n-1)}\right)(1 - d_i)\right), & p_i < p^{-i}_{(n-1)} \\ d_i \left(1 - (p_i - c_i)\sum_{k \neq i} d_k b_k f'\left(p_k - p_i\right)\right), & p_i \in (p^{-i}_{(n-1)}, p^{-i}_{(n)}) \\ d_i \left(1 - (p_i - c_i) b_i f'\left(p_i - p^{-i}_{(n)}\right)(1 - d_i)\right), & p_i > p^{-i}_{(n)} \end{cases} \tag{A-18}$$

- If $\underline{p_i < p^{-i}_{(n-1)}}$, $(p_i - c_i) f'\left(p_i - p^{-i}_{(n-1)}\right)$ is non-decreasing since $f(\cdot)$ is convex. Moreover, $1 - d_i$ is non-decreasing in $p_i$ since $\frac{\partial d_i}{\partial p_i} \leq 0$, therefore, $(p_i - c_i) b_i f'\left(p_i - p^{-i}_{(n-1)}\right)(1 - d_i)$, the product of two non-negative non-decreasing functions, is non-decreasing in $p_i$. Quasi-concavity of $\pi_i$ in $p_i$ on the interval of $[c_i, p^{-i}_{(n-1)})$ follows as in the proof of Theorem 4, see (A-14).

- If $\underline{p_i = p_{(n-1)}^{-i}}$: Clearly, $p_{(n-1)}^{-i} = p_i \geq c_i$. By (A-18),

$$\begin{aligned}
&\frac{\partial_- \pi_i}{\partial p_i}(p_{(n-1)}^{-i}, p_{-i}) - \frac{\partial_+ \pi_i}{\partial p_i}(p_{(n-1)}^{-i}, p_{-i}) \\
=\ & (p_{(n-1)}^{-i} - c_i)\left(\lim_{p_i \nearrow p_{(n-1)}^{-i}} \frac{\partial d_i}{\partial p_i}(p_i, p_{-i}) - \lim_{p_i \searrow p_{(n-1)}^{-i}} \frac{\partial d_i}{\partial p_i}(p_i, p_{-i})\right) \\
=\ & (p_{(n-1)}^{-i} - c_i)\left(\lim_{p_i \nearrow p_{(n-1)}^{-i}} \frac{\partial d_i}{\partial p_i}(p_i, p_{-i}) + \lim_{p_i \searrow p_{(n-1)}^{-i}} \sum_{j \neq i} \frac{\partial d_i}{\partial p_j}\right) \\
\geq\ & 0.
\end{aligned}$$

  The last equality follows from the fact that, by (1.24), $\frac{\partial d_i}{\partial p_i} = -d_i \sum_{j \neq i} d_j b_j f'(p_j - p_i)$ and $\frac{\partial d_i}{\partial p_j} = d_i d_j b_j f'(p_j - p_i)$ when $p_i \in (p_{(n-1)}^{-i}, p_{(n)}^{-i})$. The inequality follows from condition $(D')$. This shows that $p_{(n-1)}^{-i}$ fails to be a local minimum of the profit function $\pi_i$.

- If $\underline{p_i \in (p_{(n-1)}^{-i}, p_{(n)}^{-i})}$, by (1.24) and (A-18), $\frac{\partial \pi_i}{\partial p_i} = d_i \left(1 - (p_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial p_j}\right)$. Note that $(p_i - c_i) \sum_{j \neq i} \frac{\partial \ln(d_i)}{\partial p_j}$ is non-decreasing by $(M)$. Quasi-concavity of $\pi_i$ on the interval $(p_{(n-1)}^{-i}, p_{(n)}^{-i})$ follows, as in the proof of Theorem 4, see (A-14).

- If $\underline{p_i = p_{(n)}^{-i}}$, similar to the second case ($p_i = p_{(n-1)}^{-i}$), one shows $\frac{\partial \pi_i}{\partial p_i}(p_{(n)}^{-i}-, p_{-i}) \geq \frac{\partial \pi_i}{\partial p_i}(p_{(n)}^{-i}+, p_{-i})$ by $(D')$ so that $p_{(n)}^{-i}$ fails to be a local minimum.

- If $\underline{p_i > p_{(n)}^{-i}}$. The proof of quasi-concavity of $\pi_i$ on $(p_{(n)}^{-i}, p^{max}]$ is identical to the proof of first case, merely replacing similar $p_{(n-1)}^{-i}$ by $p_{(n)}^{-i}$.

Therefore, $\pi_i(p_i, p_{-i})$ is quasi-concave in $p_i \geq c_i$ for any $p_{-i}$. □

**Proof of Lemma 3.** Conditions $(D')$: By (1.24), we have

$$\frac{\partial d_i}{\partial p_j} = bf'(p_j - p_i)\, d_i d_j \text{ for any } p_i \in (p_{(n-1)}^{-i}, p_{(n)}^{-i}).$$

Involving the first part of (A-17), we thus obtain

$$
\begin{aligned}
& \lim_{p_i \searrow p^{-i}_{(n-1)}} \sum_{j\neq i} \frac{\partial d_i}{\partial p_j}(p_i, p_{-i}) \geq |\lim_{p_i \nearrow p^{-i}_{(n-1)}} \frac{\partial d_i}{\partial p_i}(p_i, p_{-i})| \\
\iff & \lim_{p_i \searrow p^{-i}_{(n-1)}} \sum_{j\neq i} bf'\left(p_j - p_i\right) d_i d_j \geq \lim_{p_i \nearrow p^{-i}_{(n-1)}} bf'\left(p_i - p^{-i}_{(n-1)}\right) d_i(1-d_i) \\
\overset{(e1)}{\iff} & d_i(p^{-i}_{(n-1)}, p_{-i}) \cdot \left( \sum_{j\neq i:\ p_j > p^{-i}_{(n-1)}} \alpha d_j(p^{-i}_{(n-1)}, p_{-i}) + \beta \sum_{j\neq i:\ p_j \leq p^{-i}_{(n-1)}} d_j(p^{-i}_{(n-1)}, p_{-i}) \right) \\
& \geq \beta d_i(p^{-i}_{(n-1)}, p_{-i}) \left(1 - d_i(p^{-i}_{(n-1)}, p_{-i})\right) \\
\iff & \sum_{j\neq i:\ p_j > p^{-i}_{(n-1)}} \alpha d_j(p^{-i}_{(1)}, p_{-i}) + \beta \sum_{j\neq i:\ p_j \leq p^{-i}_{(n-1)}} d_j(p^{-i}_{(n-1)}, p_{-i}) \geq \beta \left(1 - d_i(p^{-i}_{(1)}, p_{-i})\right),
\end{aligned}
\tag{A-19}
$$

where (e1) holds from the facts that $d_i$ and $d_j$ are continuous in $p_i$, $f'(x) = \alpha$ if $x > 0$ and $f'(x) = \beta$ if $x < 0$. Since $\alpha \geq \beta \geq 0$, therefore the inequality to the right of the last implication in (A-19) holds

$$
\begin{aligned}
& \sum_{j\neq i:\ p_j > p^{-i}_{(n-1)}} \alpha d_j(p^{-i}_{(n-1)}, p_{-i}) + \beta \sum_{j\neq i:\ p_j \leq p^{-i}_{(n-1)}} d_j(p^{-i}_{(n-1)}, p_{-i}) \\
& \geq \beta \sum_{j\neq i} d_j(p^{-i}_{(n-1)}, p_{-i}) = \beta \left(1 - d_i(p^{-i}_{(n-1)}, p_{-i})\right).
\end{aligned}
$$

Hence, the first inequality of $(D')$ is true. The proof of the second inequality in $(D')$ is identical to the Lemma 2.

Condition $(M)$: The proof is identical to the proof in Lemma 2. □

**Proof of** (1.26) **of Theorem 7.** We consider two cases: (i) $p_i \geq C$ (ii) $p_i < C$.

(i) $\underline{p_i \geq C}$: Let $A_i = \sum_{k\neq i: p_k \geq C} e^{a_k - b\alpha(p_k - C)}$ and $B_i = \sum_{k\neq i: p_k < C} e^{a_k - b\beta(p_k - C)}$, one has

$$
\begin{aligned}
& \left(1 - d_i^{EXO}(p_i, p_{-i})\right) - \gamma_i \left(1 - d_i^{DISC}(p_i, p_{-i})\right) \\
= \quad & \frac{e^{a_0 - b_0 f(p_0 - C)} + A_i + B_i}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i} - \gamma_i \frac{e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}}{e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}} \\
= \quad & \frac{\left(e^{a_0 - b_0 f(p_0 - C)} + A_i + B_i\right) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
& - \frac{\gamma_i \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) \cdot \left(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i\right)}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
= \quad & \frac{(1-\gamma_i)\left(e^{a_0 - b_0 f(p_0 - C)} + A_i + B_i\right) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) + \left(e^{a_0 - b_0 f(p_0 - C)} + A_i + B_i\right) \cdot e^{a_i - b\alpha\gamma_i p_i}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
& - \frac{\gamma_i \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) \cdot e^{a_i - b\alpha(p_i - C)}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
\overset{(a)}{\geq} \quad & \frac{(1-\gamma_i) e^{a_0 - b_0 f(p_0 - C)} \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) + e^{a_0 - b_0 f(p_0 - C)} \cdot e^{a_i - b\alpha\gamma_i p_i}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
& - \frac{\gamma_i \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) \cdot e^{a_i}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
= \quad & \frac{e^{a_0 - b_0 f(p_0 - C)} \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
& - \frac{\gamma_i \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) \cdot (e^{a_0 - b_0 f(p_0 - C)} + e^{a_i})}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\alpha(p_i - C)} + A_i + B_i) \cdot \left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
\overset{(b)}{\geq} \quad & 0, \text{ thus verifying (1.26).}
\end{aligned}
$$

Here, inequality (a) follows from the fact that $\gamma_i \leq 1$ and $A_i, B_i \geq 0$; inequality (b) follows from the fact that

$$
\gamma_i \leq \frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i}} \leq \frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i}} \cdot \frac{e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}}{e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}} \quad \text{for any } p,
$$

where the first inequality follows from

$$
\begin{aligned}
& \delta_i = 1 - \gamma_i \geq 1 - \frac{\beta}{\alpha} \cdot \frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i + b\beta(C - c_i)}} \\
\Longleftrightarrow \quad & \gamma_i \leq \frac{\beta}{\alpha} \cdot \frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i + b\beta(C - c_i)}} \leq \frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i}}.
\end{aligned}
$$

(ii) $\underline{p_i < C}$: One has

$$
\begin{aligned}
& \frac{\beta}{\alpha}\left(1 - d_i^{EXO}(p_i, p_{-i})\right) - \gamma_i\left(1 - d_i^{DISC}(p_i, p_{-i})\right) \\
= \; & \frac{\beta}{\alpha}\frac{e^{a_0 - b_0 f(p_0 - C)} + A_i + B_i}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(p_i - C)} + A_i + B_i} - \gamma_i \frac{e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}}{e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}} \\
= \; & \frac{(\frac{\beta}{\alpha} - \gamma_i)\left(e^{a_0 - b_0 f(p_0 - C)} + A_i + B_i\right)\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) + \frac{\beta}{\alpha}\left(e^{a_0 - b_0 f(p_0 - C)} + A_i + B_i\right)\cdot e^{a_i - b\alpha\gamma_i p_i}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(p_i - C)} + A_i + B_i)\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
& - \frac{\gamma_i\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)\cdot e^{a_i - b\beta(p_i - C)}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(p_i - C)} + A_i + B_i)\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
\overset{(c)}{\geq} \; & \frac{(\frac{\beta}{\alpha} - \gamma_i)e^{a_0 - b_0 f(p_0 - C)}\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right) + \frac{\beta}{\alpha}e^{a_0 - b_0 f(p_0 - C)}\cdot e^{a_i - b\alpha\gamma_i p_i}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(p_i - C)} + A_i + B_i)\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
& - \frac{\gamma_i\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)\cdot e^{a_i - b\beta(c_i - C)}}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(p_i - C)} + A_i + B_i)\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
= \; & \frac{\frac{\beta}{\alpha}e^{a_0 - b_0 f(p_0 - C)}\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(p_i - C)} + A_i + B_i)\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
& - \frac{\gamma_i\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)\cdot(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(c_i - C)})}{(e^{a_0 - b_0 f(p_0 - C)} + e^{a_i - b\beta(p_i - C)} + A_i + B_i)\cdot\left(e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}\right)} \\
\overset{(d)}{\geq} \; & 0,
\end{aligned}
$$

where inequality (c) follows from the fact that $p_i \geq c_i$, $\gamma_i = 1 - \delta_i \leq \beta/\alpha$ and $A_i, B_i \geq 0$; inequality (d) follows from the fact that

$$
\begin{aligned}
\gamma_i & \leq \frac{\beta}{\alpha}\cdot\frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i + b\beta(C - c_i)}} \\
& \leq \frac{\beta}{\alpha}\cdot\frac{e^{a_0 - b_0 f(p_0 - C)}}{e^{a_0 - b_0 f(p_0 - C)} + e^{a_i + b\beta(C - c_i)}}\cdot\frac{e^{a_0 - b_0\alpha\gamma_0 p_0} + e^{a_i - b\alpha\gamma_i p_i} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}}{e^{a_0 - b_0\alpha\gamma_0 p_0} + \sum_{k\neq i} e^{a_k - b\alpha\gamma_k p_k}}.
\end{aligned}
$$

□

**Proof of Corollary 1.** We prove the first threshold result. Consider

$$
\begin{aligned}
& C \leq \min_i \left\{ c_i + (b\beta)^{-1}\left(a_0 - a_i + \log\left(\frac{\beta}{\alpha(1-\delta)} - 1\right)\right)\right\} \\
\iff \; & C \leq c_i + (b\beta)^{-1}\left(a_0 - a_i + \log\left(\frac{\beta}{\alpha(1-\delta)} - 1\right)\right), \quad \forall i \\
\iff \; & a_i - a_0 + (b\beta)(C - c_i) \leq \log\left(\frac{\beta}{\alpha(1-\delta)} - 1\right), \quad \forall i \\
\iff \; & \frac{\beta}{\alpha}\left(\frac{e^{a_0}}{e^{a_0} + e^{a_i - b\beta(C - c_i)}}\right) \leq 1 - \delta, \quad \forall i.
\end{aligned}
$$

It then follows from Theorem 7(b) that $p_i^{*DISC} \geq \overline{p_i}^{*EXO}$. The threshold result follows from the monotonicity of $\overline{p_i}^{*EXO}$ in $C$, see Proposition 2. □

**Proof of Theorem 8.**

(a) Without loss of generality, we set $\alpha = 1$ in the proof. Since the constant reference value $C \leq c_{min}$, $C \leq p$ for any feasible price vector $p \in [p^{min}, p^{max}]$. Therefore, one has

$$\begin{aligned} d_i^{EXO}(\mathbf{p}) &= \frac{\exp\left(a_i - b(p_i - C)\right)}{e^{a_0} + \exp\left(a_i - b(p_i - C)\right) + \sum_{k \neq i} \exp\left(a_k - b(p_k - C)\right)} \\ &= \frac{e^{a_i - bp_i}}{e^{a_0 - bC} + e^{a_i - bp_i} + \sum_{k \neq i} e^{a_k - bp_k}}. \end{aligned}$$

Similarly, one obtains the market share under the lowest reference value as follows:

$$\begin{aligned} d_i^{LOW}(\mathbf{p}) &= \begin{cases} \frac{\exp\left(a_i - b(p_i - p_{(1)}^{-i})\right)}{e^{a_0} + \exp\left(a_i - b(p_i - p_{(1)}^{-i})\right) + \sum_{k \neq i} \exp\left(a_k - b(p_k - p_{(1)}^{-i})\right)}, & p_i \geq p_{(1)}^{-i} \\ \frac{\exp(a_i)}{e^{a_0} + \exp(a_i) + \sum_{k \neq i} \exp(a_k - b\alpha(p_k - p_i))}, & p_i < p_{(1)}^{-i} \end{cases} \\ &= \begin{cases} \frac{e^{a_i - bp_i}}{e^{a_0 - bp_{(1)}^{-i}} + e^{a_i - bp_i} + \sum_{k \neq i} e^{a_k - bp_k}}, & p_i \geq p_{(1)}^{-i} \\ \frac{e^{a_i - bp_i}}{e^{a_0 - bp_i} + e^{a_i - bp_i} + \sum_{k \neq i} e^{a_k - bp_k}}, & p_i < p_{(1)}^{-i} \end{cases}. \end{aligned}$$

Since $C \leq p^{min}$, it follows that, $d_i^{EXO}(\mathbf{p}) \leq d_i^{LOW}(\mathbf{p})$ for any feasible price vector $\mathbf{p}$. Again, since $C \leq p^{min}$, by (1.11), one has

$$\begin{aligned} \frac{\partial_+ \log \pi_i^{EXO}(p_i, p_{-i})}{\partial p_i} &= \frac{1}{p_i - c_i} - bf'_+(p_i - C)\left(1 - d_i^{EXO}(p_i, p_{-i})\right) \\ &= \frac{1}{p_i - c_i} - b\left(1 - d_i^{EXO}(p_i, p_{-i})\right), \end{aligned}$$

By (1.19) and (1.21), one has

$$\frac{\partial_+ \log \pi_i^{LOW}(p_i, p_{-i})}{\partial p_i} = \begin{cases} b\left(1 - d_i^{LOW}(p_i, p_{-i})\right), & p_i \geq p_{(1)}^{-i} \\ b\left(1 - d_i^{LOW}(p_i, p_{-i}) - d_0^{LOW}(p_i, p_{-i})\right), & p_i < p_{(1)}^{-i} \end{cases}$$

Therefore, one obtains

$$\begin{aligned} &\frac{\partial_+ \log \pi_i^{EXO}(p_i, p_{-i})}{\partial p_i} - \frac{\partial_+ \log \pi_i^{LOW}(p_i, p_{-i})}{\partial p_i} \\ =& \begin{cases} b\left(d_i^{EXO}(p_i, p_{-i}) - d_i^{LOW}(p_i, p_{-i})\right), & p_i \geq p_{(1)}^{-i} \\ b\left(d_i^{EXO}(p_i, p_{-i}) - d_i^{LOW}(p_i, p_{-i}) - d_0^{LOW}(p_i, p_{-i})\right), & p_i < p_{(1)}^{-i} \end{cases} \\ \leq& \; 0. \quad \text{since } d_i^{EXO}(\mathbf{p}) \leq d_i^{LOW}(\mathbf{p}) \text{ for any feasible } p. \end{aligned}$$

Therefore, the best response functions satisfy $p_i^{*LOW}(p_{-i}) \geq p_i^{*EXO}(p_{-i})$ for any firm $i$ and any $p_{-i}$. The remainder of the proof is analogous to that of Theorem 7 (b).

(b) Since $d_0 \geq 1 - \beta/\alpha$, it follows from Proposition 2 that both the component-wise smallest and largest equilibrium $\underline{p}^{*EXO}$ and $\overline{p}^{*EXO}$ are monotonically increasing in $C$. The threshold result follows immediately.

□

**Proof of Theorem 9.** (a) By (1.11), one has

$$\begin{aligned}\frac{\partial_+ \log \pi_i^{EXO}(p_i, p_{-i})}{\partial p_i} &= \frac{1}{p_i - c_i} - bf'_+(p_i - C)\left(1 - d_i^{EXO}(p_i, p_{-i})\right) \\ &= \begin{cases} \frac{1}{p_i - c_i} - b\alpha\left(1 - d_i^{EXO}(p_i, p_{-i})\right), & p_i \geq C \\ \frac{1}{p_i - c_i} - b\beta\left(1 - d_i^{EXO}(p_i, p_{-i})\right), & p_i < C \end{cases}\end{aligned}$$

In the (PSOG) case, (1.19) and (1.21) coincide, so that

$$\frac{\partial_+ \log \pi_i^{LOW}(p_i, p_{-i})}{\partial p_i} = \frac{1}{p_i - c_i} - b\alpha\left(1 - d_i^{LOW}(p_i, p_{-i})\right).$$

Thus, we obtain

$$\begin{aligned}&\frac{\partial_+ \log \pi_i^{EXO}(p_i, p_{-i})}{\partial p_i} - \frac{\partial_+ \log \pi_i^{LOW}(p_i, p_{-i})}{\partial p_i} \\ &= \begin{cases} b\alpha\left(d_i^{EXO}(p_i, p_{-i}) - d_i^{LOW}(p_i, p_{-i})\right), & p_i \geq C \\ b\alpha\left(1 - d_i^{LOW}(p_i, p_{-i})\right) - b\beta\left(1 - d_i^{EXO}(p_i, p_{-i})\right), & p_i < C \end{cases}. \qquad \text{(A-20)}\end{aligned}$$

In the Appendix, we prove that, for $C \leq c_{min} + \frac{\ln(\alpha) - \ln(\beta)}{b(\alpha - \beta)}$, the expression in (A-20) is non-negative. Recall that, both $\log \pi_i^{LOW}(p_i, p_{-i})$ and $\log \pi_i^{EXO}(p_i, p_{-i})$ are quasi-concave by Theorem 4 and Proposition 1, respectively. The non-negativity of (A-20) thus implies

$$p_i^{*LOW}(p_{-i}) \equiv \sup\left\{p_i : \frac{\partial_+ \log \pi_i^{LOW}(p_i, p_{-i})}{\partial p_i} \geq 0\right\} \leq \sup\left\{p_i : \frac{\partial_+ \log \pi_i^{EXO}(p_i, p_{-i})}{\partial p_i} \geq 0\right\} \equiv p_i^{*EXO}(p_{-i}).$$

The inequality follows from the fact that the set to the right contains the set to the left, see (A-20). The remainder of the proof is analogous to that of Theorem 7 (b).

(b) Follows immediately from part (a) and the monotonicity of $\overline{p}^{*EXO}(C)$ and $\underline{p}^{*EXO}(C)$ in $C$, see Theorem 2. □

**Proof of** (A-20) **in Theorem 9.** Note by (1.20) that the sales volume in the lowest price subsidy model satisfies

$$d_i^{LOW}(\mathbf{p}) = \frac{\exp(a_i - b\alpha p_i)}{\exp(a_i - b\alpha p_i) + \sum_{k\neq i}\exp(a_k - b\alpha p_k)}.$$

We distinguish between the following two cases: (i) $p_{(1)}^{-i} \geq C$; (ii) $p_{(1)}^{-i} < C$.

(i) $\underline{p_{(1)}^{-i} \geq C}$. By (1.10), since $p_k \geq p_{(1)}^{-i} \geq C$ for any $k \neq i$, one has

$$\begin{aligned} d_i^{EXO}(\mathbf{p}) &= \begin{cases} \frac{\exp(a_i - b\alpha(p_i - C))}{\exp(a_i - b\alpha(p_i - C)) + \sum_{k\neq i}\exp(a_k - b\alpha(p_k - C))}, & p_i \geq C \\ \frac{\exp(a_i - b\beta(p_i - C))}{\exp(a_i - b\beta(p_i - C)) + \sum_{k\neq i}\exp(a_k - b\alpha(p_k - C))}, & p_i < C \end{cases} \\ &= \begin{cases} \frac{\exp(a_i - b\alpha p_i)}{\exp(a_i - b\alpha p_i) + \sum_{k\neq i}\exp(a_k - b\alpha p_k)}, & p_i \geq C \\ \frac{\exp(a_i - b(\beta p_i + (\alpha-\beta)C))}{\exp(a_i - b(\beta p_i + (\alpha-\beta)C)) + \sum_{k\neq i}\exp(a_k - b\alpha p_k)}, & p_i < C \end{cases}. \end{aligned}$$

It is obvious that $d_i^{EXO}(\mathbf{p}) = d_i^{LOW}(\mathbf{p})$ when $p_i \geq C$, thus, (A-20) holds. When $p_i < C$, $p_{(1)} = p_i$, and

$$\begin{aligned} & \alpha\left(1 - d_i^{LOW}(\mathbf{p})\right) - \beta\left(1 - d_i^{EXO}(\mathbf{p})\right) \\ = \ & \alpha\frac{\sum_{k\neq i}\exp(a_k - b\alpha p_k)}{\exp(a_i - b\alpha p_i) + \sum_{k\neq i}\exp(a_k - b\alpha p_k)} \qquad \text{(A-21)} \\ & -\beta\frac{\sum_{k\neq i}\exp(a_k - b\alpha p_k)}{\exp(a_i - b(\beta p_i + (\alpha-\beta)C)) + \sum_{k\neq i}\exp(a_k - b\alpha p_k)} \\ = \ & \frac{\left(\sum_{k\neq i} e^{a_k - b\alpha p_k}\right)\cdot\left[(\alpha-\beta)\cdot\left(\sum_{k\neq i} e^{a_k - b\alpha p_k}\right) + \alpha e^{a_i - b(\beta p_i + (\alpha-\beta)C)} - \beta e^{a_i - b\alpha p_i}\right]}{\left[e^{a_i - b\alpha p_i} + \sum_{k\neq i} e^{a_k - b\alpha p_k}\right]\cdot\left[e^{a_i - b(\beta p_i + (\alpha-\beta)C)} + \sum_{k\neq i} e^{a_k - b\alpha p_k}\right]} \\ = \ & \frac{\left(\sum_{k\neq i} e^{a_k - b\alpha p_k}\right)\cdot\left[(\alpha-\beta)\cdot\left(\sum_{k\neq i} e^{a_k - b\alpha p_k}\right) + \alpha e^{a_i - b\alpha p_i}\left(e^{-b(\alpha-\beta)(C-p_i)} - \frac{\beta}{\alpha}\right)\right]}{\left[e^{a_i - b\alpha p_i} + \sum_{k\neq i} e^{a_k - b\alpha p_k}\right]\cdot\left[e^{a_i - b(\beta p_i + (\alpha-\beta)C)} + \sum_{k\neq i} e^{a_k - b\alpha p_k}\right]} \qquad \text{(A-22)} \\ \geq \ & \frac{\left(\sum_{k\neq i} e^{a_k - b\alpha p_k}\right)\cdot\left[(\alpha-\beta)\cdot\left(\sum_{k\neq i} e^{a_k - b\alpha p_k}\right) + \alpha e^{a_i - b\alpha p_i}\left(e^{-b(\alpha-\beta)(C-c_i)} - \frac{\beta}{\alpha}\right)\right]}{\left[e^{a_i - b\alpha p_i} + \sum_{k\neq i} e^{a_k - b\alpha p_k}\right]\cdot\left[e^{a_i - b(\beta p_i + (\alpha-\beta)C)} + \sum_{k\neq i} e^{a_k - b\alpha p_k}\right]} \\ \geq \ & 0, \end{aligned}$$

where the first inequality follows from $p_i \geq c_i$, and the second inequality from $\alpha \geq \beta$ and $e^{-b(\alpha-\beta)(C-c_i)} \geq \frac{\beta}{\alpha}$ since $C \leq c_{min} + \frac{\ln(\alpha) - \ln(\beta)}{b(\alpha-\beta)}$.

(ii) $\underline{p_{(1)}^{-i} < C}$. By (1.10), one has

$$
\begin{aligned}
d_i^{EXO}(\mathbf{p}) &= \begin{cases} \frac{\exp(a_i - b\alpha(p_i - C))}{\exp(a_i - b\alpha(p_i - C)) + \sum_{k\neq i:\ p_k<C} \exp(a_k - b\beta(p_k - C)) + \sum_{k\neq i:\ p_k\geq C} \exp(a_k - b\alpha(p_k - C))}, & p_i \geq C \\ \frac{\exp(a_i - b\beta(p_i - C))}{\exp(a_i - b\beta(p_i - C)) + \sum_{k\neq i:\ p_k<C} \exp(a_k - b\beta(p_k - C)) + \sum_{k\neq i:\ p_k\geq C} \exp(a_k - b\alpha(p_k - C))}, & p_i < C \end{cases} \\
&\geq \begin{cases} \frac{\exp(a_i - b\alpha(p_i - C))}{\exp(a_i - b\alpha(p_i - C)) + \sum_{k\neq i:\ p_k<C} \exp(a_k - b\alpha(p_k - C)) + \sum_{k\neq i:\ p_k\geq C} \exp(a_k - b\alpha(p_k - C))}, & p_i \geq C \\ \frac{\exp(a_i - b\beta(p_i - C))}{\exp(a_i - b\beta(p_i - C)) + \sum_{k\neq i:\ p_k<C} \exp(a_k - b\alpha(p_k - C)) + \sum_{k\neq i:\ p_k\geq C} \exp(a_k - b\alpha(p_k - C))}, & p_i < C \end{cases} \\
&= \begin{cases} \frac{\exp(a_i - b\alpha p_i)}{\exp(a_i - b\alpha p_i) + \sum_{k\neq i} \exp(a_k - b\alpha p_k)} = d_i^{LOW}(\mathbf{p}), & p_i \geq C \\ \frac{\exp(a_i - b(\beta p_i + (\alpha-\beta)C))}{\exp(a_i - b(\beta p_i + (\alpha-\beta)C)) + \sum_{k\neq i} \exp(a_k - b\alpha p_k)}, & p_i < C \end{cases},
\end{aligned}
$$

where the inequality follows by replacing, in the denominator, each of the term in the index set $\{k \neq i, p_k < C\}$ by a larger value, since $0 < \beta \leq \alpha$. This proves (A-20) when $p_i \geq C$. When $p_i < C$, one has

$$
\begin{aligned}
& \alpha\left(1 - d_i^{LOW}(\mathbf{p})\right) - \beta\left(1 - d_i^{EXO}(\mathbf{p})\right) \\
\geq\ & \alpha \frac{\sum_{k\neq i} \exp(a_k - b\alpha p_k)}{\exp(a_i - b\alpha p_i) + \sum_{k\neq i} \exp(a_k - b\alpha p_k)} \\
& -\beta \frac{\sum_{k\neq i} \exp(a_k - b\alpha p_k)}{\exp(a_i - b(\beta p_i + (\alpha-\beta)C)) + \sum_{k\neq i} \exp(a_k - b\alpha p_k)} \\
\geq\ & 0. \quad \text{by (A-22), ( Note that the left hand side coincides with the second expression in (A-22). )}
\end{aligned}
$$

We have thus shown that the inequalities (A-20) apply for any firm $i$ and any $\mathbf{p}$. □

# Appendix B

# Appendix for Chapter 2

## B.1 Proofs

In this section, we provide a structural estimation methodology to estimate the parameters in the general competition model with switching cost. For any $i = 1, 2, \ldots, N$, let

$$\Delta_i = a_i - b_i f(p_i - C) - a_0 + b_0 f(p_0 - C),$$

Thus, determining the intercepts $\{a_i\}$ is equivalent to computing the quantities $\{\Delta_i\}_{i=1}^N$. (Recall $a_0 = 0$, by normalization.) We have, by (2.10)–(2.13), for any MA plan $i = 1, 2, \ldots, N$, since we normalize $S_2 = 0$, that

$$\begin{aligned} d_i &= h_1 \frac{e^{a_i - b_i \cdot f(p_i - g(\mathbf{p}))}}{e^{S_1} \cdot e^{a_0 - b_0 \cdot f(p_0 - g(\mathbf{p}))} + \sum_{k=1}^N e^{a_k - b_k \cdot f(p_k - g(\mathbf{p}))}} \\ &\quad + h_2 \frac{e^{a_i - b_i \cdot f(p_i - g(\mathbf{p}))}}{e^{a_0 - b_0 \cdot f(p_0 - g(\mathbf{p}))} + \sum_{k=1}^N e^{a_k - b_k \cdot f(p_k - g(\mathbf{p}))}} \\ &= h_1 \frac{e^{\Delta_i}}{e^{S_1} + \sum_{k=1}^N e^{\Delta_k}} + h_2 \frac{e^{\Delta_i}}{1 + \sum_{k=1}^N e^{\Delta_k}} \\ &= e^{\Delta_i} \cdot \left[ \frac{h_1}{e^{S_1} + \sum_{k=1}^N e^{\Delta_k}} + \frac{h_2}{1 + \sum_{k=1}^N e^{\Delta_k}} \right]. \end{aligned}$$

Similarly, the sales volume of the FFS plan may be represented as

$$d_0 = \frac{h_1 e^{S_1}}{e^{S_1} + \sum_{k=1}^N e^{\Delta_k}} + \frac{h_2}{1 + \sum_{k=1}^N e^{\Delta_k}}.$$

Dividing the sales volume of MA plan $i, i \geq 2$ by the sales volume of MA plan 1, we obtain

$$\ln\left(\frac{d_i}{d_1}\right) = \Delta_i - \Delta_1 \quad \text{or} \quad \Delta_i = \Delta_1 + \ln(d_i) - \ln(d_1), \quad i = 1, 2, \ldots, N. \tag{B-1}$$

Thus, all $\{\Delta_i\}$ are specified once $\Delta_1$ is computed. Dividing the market share of MA plan 1 by the market share of the FFS plan 0, we also have, by (2.10)–(2.13),

$$\begin{aligned}
\frac{d_1}{d_0} &= e^{\Delta_1} \cdot \frac{h_1\left(1+\sum_{k=1}^N e^{\Delta_k}\right) + h_2\left(e^{S_1}+\sum_{k=1}^N e^{\Delta_k}\right)}{h_1 e^{S_1}\left(1+\sum_{k=1}^N e^{\Delta_k}\right) + h_2\left(e^{S_1}+\sum_{k=1}^N e^{\Delta_k}\right)} \\
&= e^{\Delta_1} \cdot \frac{(h_1+h_2e^{S_1}) + (h_1+h_2)\sum_{k=1}^N e^{\Delta_k}}{(h_1+h_2)e^{S_1} + (h_2+h_1e^{S_1})\sum_{k=1}^N e^{\Delta_k}} \\
&= e^{\Delta_1} \cdot \frac{(h_1+h_2e^{S_1}) + (h_1+h_2)e^{\Delta_1}\sum_{k=1}^N e^{\Delta_k-\Delta_1}}{(h_1+h_2)e^{S_1} + (h_2+h_1e^{S_1})e^{\Delta_1}\sum_{k=1}^N e^{\Delta_k-\Delta_1}} \\
&= e^{\Delta_1} \cdot \frac{(h_1+h_2e^{S_1}) + (h_1+h_2)e^{\Delta_1}\sum_{k=1}^N d_k/d_1}{(h_1+h_2)e^{S_1} + (h_2+h_1e^{S_1})e^{\Delta_1}\sum_{k=1}^N d_k/d_1}.
\end{aligned}$$

Recall that the sizes of the two segments add up to 1, i.e., $h_1 + h_2 = 1$, thus, $\Delta_1$ satisfies

$$\begin{aligned}
0 &= \frac{\sum_{k=1}^N d_k}{d_1}(e^{\Delta_1})^2 + \left(h_1 + h_2e^{S_1} - (h_2+h_1e^{S_1})\frac{\sum_{k=1}^N d_k}{d_0}\right)e^{\Delta_1} - \frac{d_1}{d_0}e^{S_1} \\
&= \frac{1-d_0}{d_1}(e^{\Delta_1})^2 + \left(h_1 + h_2e^{S_1} - (h_2+h_1e^{S_1})\frac{1-d_0}{d_0}\right)e^{\Delta_1} - \frac{d_1}{d_0}e^{S_1}.
\end{aligned}$$

The unique positive root of this quadratic equation in $e^{\Delta_1}$ is:

$$\begin{aligned}
e^{\Delta_1} &= \frac{-\left(h_1+h_2e^{S_1} - (h_2+h_1e^{S_1})\frac{1-d_0}{d_0}\right) + \sqrt{\left(h_1+h_2e^{S_1} - (h_2+h_1e^{S_1})\frac{1-d_0}{d_0}\right)^2 + 4\frac{1-d_0}{d_0}e^{S_1}}}{2(1-d_0)/d_1} \\
&= \frac{d_1}{d_0} \cdot \frac{-\hat{B} + \sqrt{\hat{B}^2 + 4d_0(1-d_0)e^{S_1}}}{2(1-d_0)}. \qquad \text{(B-2)}
\end{aligned}$$

Here

$$\hat{B} = (h_1+h_2e^{S_1})d_0 - (h_2+h_1e^{S_1})(1-d_0)$$

Thus, by (B-1)–(B-2), for any $i = 1, 2, \ldots, N$, we have

$$\Delta_i = \ln(d_i) - \ln(d_0) + \ln\left(\frac{-\hat{B} + \sqrt{\hat{B}^2 + 4d_0(1-d_0)e^{S_1}}}{2(1-d_0)}\right). \qquad \text{(B-3)}$$

With all parameters of the demand functions fully specified, the remaining task is to identify the marginal cost rates $\{c_i\}$. As with the models in Sections 2.4, this is accomplished by observing that the observed price vector $p^*$ is an interior point of the feasible price space and by assuming that it is a Nash equilibrium. It is easily verified, as in Section 2.4, that each profit function $\pi_i$ is differentiable in its own price $p_i$, unless $p_i = C$, in which case the right and left derivatives $\frac{\partial_+ \pi_i}{\partial p_i}$ and $\frac{\partial_- \pi_i}{\partial p_i}$ exist. This implies that

$$0 = \frac{\partial \pi_i(p^*)}{\partial p_i} = d_i(p^*) + (p_i^* - c_i)\frac{\partial d_i}{\partial p_i}; \quad \text{if } p_i^* \neq C. \tag{B-4}$$

$$0 \geq \frac{\partial_+ \pi_i(p^*)}{\partial p_i} = d_i(p^*) + (C - c_i)\frac{\partial_+ d_i}{\partial p_i}; \quad \text{if } p_i^* = C, \tag{B-5}$$

$$0 \leq \frac{\partial_- \pi_i(p^*)}{\partial p_i} = d_i(p^*) + (C - c_i)\frac{\partial_- d_i}{\partial p_i}; \quad \text{if } p_i^* = C.$$

Thus, as in the model of Section 2.4, if $p_i^* \neq C$, the marginal cost rate $c_i$ is determined as the unique root of equation (B-4). If $p_i^* = C$, (B-5) determines an interval for the cost rate $c_i$.

## B.2 Additional Results

Table B.1: Price equilibrium ($/month) for exogenous subsidy v.s. endogenous subsidy

| | No inertia cost | | | With inertia cost | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $S_1 = 0$ | | | $S_1 = 2$ | | | $S_1 = 4$ | | |
| | Exog. | Lowest | Second | Exog. | Lowest | Second | Exog. | Lowest | Second |
| FFS | 752.96 | 746.92 | 747.36 | 752.96 | 747.89 | 748.28 | 752.96 | 749.67 | 749.94 |
| 1st MA plan | 669.71 | 641.03 | 657.96 | 669.71 | 642.98 | 659.72 | 669.71 | 643.51 | 661.85 |
| 2nd MA plan | 706.26 | 677.70 | 679.21 | 706.26 | 673.45 | 674.17 | 706.26 | 678.31 | 677.67 |
| 3rd MA plan | 727.98 | 703.97 | 704.03 | 727.98 | 705.87 | 706.07 | 727.98 | 707.52 | 706.36 |
| capitation | 837.68 | 640.64 | 683.32 | 837.68 | 641.48 | 684.15 | 837.68 | 644.04 | 686.61 |
| ave. price | 752.82 | 740.95 | 742.47 | 752.82 | 742.75 | 744.14 | 752.82 | 746.33 | 747.43 |

Note1: The column headings “Exog.”, “Lowest” or “Second” are short hand for payment schemes in which the capitation rate is set exogenously, or determined as the lowest or the second-lowest bid, respectively.

Note2: The row headings “1st MA plan”, “2nd MA plan”, “3rd MA plan” are short hand for the MA plan with the lowest, second-lowest or third-lowest bid, respectively.

Table B.2: Market share (%) for exogenous subsidy v.s. endogenous subsidy

| | No inertia cost | | | With inertia cost | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $S_1 = 0$ | | | $S_1 = 2$ | | | $S_1 = 4$ | | |
| | Exog. | Lowest | Second | Exog. | Lowest | Second | Exog. | Lowest | Second |
| FFS | 78.05 | 72.88 | 73.25 | 78.05 | 74.26 | 74.55 | 78.05 | 76.52 | 76.65 |
| 1st MA plan | 2.58 | 3.77 | 3.17 | 2.58 | 3.53 | 2.88 | 2.58 | 2.76 | 2.23 |
| 2nd MA plan | 2.34 | 3.23 | 3.19 | 2.34 | 2.68 | 2.72 | 2.34 | 2.43 | 2.52 |
| 3rd MA plan | 2.63 | 3.20 | 3.25 | 2.63 | 3.21 | 3.28 | 2.63 | 2.19 | 2.19 |
| 4th MA plan | 2.25 | 3.19 | 3.20 | 2.25 | 2.87 | 2.93 | 2.25 | 2.95 | 3.00 |
| 5th MA plan | 2.34 | 2.86 | 2.93 | 2.34 | 2.71 | 2.71 | 2.34 | 2.44 | 2.37 |

Table B.3: Distribution of out-of-pocket payment for FFS and MA enrollees

| out-of-pocket payment ($/month) | No inertia cost | | | With inertia cost | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | $S_1 = 0$ | | | $S_1 = 2$ | | | $S_1 = 4$ | | |
| | Exog. | Lowest | Second | Exog. | Lowest | Second | Exog. | Lowest | Second |
| 0 | 89.12% | 16.49% | 28.71%* | 89.12% | 16.62% | 28.54% | 89.12% | 17.34% | 28.89% |
| (0,40] | 8.07% | 16.35% | 21.48% | 8.07% | 16.13% | 21.65% | 8.07% | 16.03% | 21.21% |
| (40, 70] | 2.03% | 14.61% | 14.68% | 2.03% | 14.45% | 14.35% | 2.03% | 14.15% | 14.47% |
| (70,125] | 0.68% | 23.50% | 19.12% | 0.68% | 23.10% | 19.01% | 0.68% | 22.52% | 17.56% |
| (125,200] | 0.07% | 14.91% | 9.65% | 0.07% | 15.32% | 9.79% | 0.07% | 14.93% | 10.70% |
| > 200 | 0.03% | 14.13% | 6.36% | 0.03% | 14.38% | 6.66% | 0.03% | 15.04% | 7.17% |
| Ave. ($/month) | 3.37 | 100.31 | 64.25 | 3.37 | 101.26 | 65.28 | 3.37 | 102.29 | 66.83 |

Note*: This number (28.71%) is much larger than the sum of market shares for the lowest MA plan (3.17%) and second-lowest MA plan (3.19%). The difference averages from the beneficiaries enrolled in FFS plan, whose price is less than the subsidy.

Table B.4: Estimation Result via Estimation (Plan II and inertia cost $S = 2$)

| | | Estimation Model | | | Calibration Model | | |
|---|---|---|---|---|---|---|---|
| | Exogenous | Lowest | Second | Exogenous ($C - 50$) | Lowest | Second | Exogenous ($C - 50$) |
| **Prices** | | | | | | | |
| FFS | 752.96 | 743.97 | 744.15 | 747.17 | 747.89 | 748.28 | 752.57 |
| Capitation | 837.68 | 602.07 | 644.32 | 787.68 | 641.48 | 684.15 | 787.68 |
| ave. price | 753.82 | 747.58 | 747.74 | 750.30 | 742.75 | 744.14 | 753.48 |
| **Market Share** | | | | | | | |
| FFS | 78.05% | 91.83% | 92.12% | 94.59% | 74.26% | 74.55% | 78.07% |
| 1st MA plan | 2.58% | 0.61% | 0.57% | 0.40% | 3.53% | 2.88% | 2.46% |
| 2nd MA plan | 2.34% | 0.60% | 0.57% | 0.40% | 2.68% | 2.72% | 2.05% |
| 3rd MA plan | 2.63% | 0.61% | 0.57% | 0.40% | 3.21% | 3.28% | 2.41% |
| **Out-of-pocket payment per month** | | | | | | | |
| 0 | 89.12% | 1.52% | 11.46% | 69.88% | 16.62% | 28.54% | 69.64% |
| (0,40] | 8.07% | 12.59% | 18.34% | 16.68% | 16.13% | 21.65% | 15.80% |
| (40, 70] | 2.03% | 12.06% | 14.36% | 8.40% | 14.45% | 14.35% | 8.16% |
| (70,125] | 0.68% | 23.10% | 19.32% | 4.61% | 23.10% | 19.01% | 5.69% |
| (125,200] | 0.07% | 22.57% | 17.25% | 0.43% | 15.32% | 9.79% | 0.67% |
| $> 200$ | 0.03% | 28.16% | 19.27% | 0.01% | 14.38% | 6.66% | 0.04% |
| ave. payment | 3.37 | 145.51 | 106.73 | 12.09 | 101.26 | 65.28 | 13.23 |
| **Government payment ($) per month** | | | | | | | |
| payment | 815.87 | 599.96 | 641.82 | 775.75 | 641.80 | 683.50 776.12 | |
| % of Exogenous | | 73.54% | 78.67% | 95.08% | 78.66% | 83.78% | 95.13% |

Table B.5: Estimation Result via Estimation (Plan I and inertia cost $S = 2$)

| | | Estimation Model | | | Calibration Model | | |
|---|---|---|---|---|---|---|---|
| | Exogenous | Lowest | Second | Exogenous ($C - 50$) | Lowest | Second | Exogenous ($C - 50$) |
| **Prices** | | | | | | | |
| FFS | 752.96 | 754.72 | 754.66 | 747.98 | 756.52 | 757.11 | 753.37 |
| Capitation | 837.68 | 599.51 | 642.26 | 787.68 | 639.11 | 682.10 | 787.68 |
| ave. price | 753.82 | 754.81 | 754.83 | 751.75 | 743.49 | 746.27 | 756.84 |
| **Market Share** | | | | | | | |
| FFS | 78.05% | 99.26% | 98.86% | 93.46% | 94.43% | 92.28% | 83.87% |
| 1st MA plan | 2.58% | 0.07% | 0.11% | 0.49% | 1.09% | 1.25% | 1.86% |
| 2nd MA plan | 2.34% | 0.07% | 0.11% | 0.49% | 0.90% | 1.27% | 1.56% |
| 3rd MA plan | 2.63% | 0.07% | 0.11% | 0.49% | 0.77% | 1.08% | 2.05% |
| **Out-of-pocket payment per month** | | | | | | | |
| 0 | 97.38% | 95.27% | 94.47% | 97.12% | 99.26% | 98.86% | 98.67% |
| (0,40] | 1.82% | 1.37% | 2.24% | 1.30% | 0.01% | 0.07% | 0.74% |
| (40, 70] | 0.45% | 0.94% | 1.24% | 0.72% | 0.03% | 0.18% | 0.58% |
| (70,125] | 0.26% | 1.64% | 1.53% | 0.68% | 0.24% | 0.27% | 0 |
| (125,200] | 0.06% | 0.70% | 0.48% | 0.17% | 0.13% | 0.36% | 0.00% |
| $> 200$ | 0.03% | 0.09% | 0.05% | 0.02% | 0.34% | 0.26% | 0.00% |
| ave. payment | 0.92 | 3.63 | 3.40 | 1.50 | 1.29 | 1.53 | 0.42 |
| **Government payment ($) per month** | | | | | | | |
| payment | 815.87 | 753.51 | 753.30 | 756.13 | 750.67 | 751.31 | 759.89 |
| % of Exogenous | | 92.36% | 92.33% | 92.68% | 92.01% | 92.09% | 93.14% |

Table B.6: Estimation Result via Estimation (Plan II and inertia cost $S = 4$)

| | | Estimation Model | | | Calibration Model | | |
|---|---|---|---|---|---|---|---|
| | Exogenous | Lowest | Second | Exogenous ($C - 50$) | Lowest | Second | Exogenous ($C - 50$) |
| **Prices** | | | | | | | |
| FFS | 752.96 | 747.25 | 747.33 | 749.66 | 749.67 | 749.94 | 752.77 |
| Capitation | 837.68 | 602.71 | 644.88 | 787.68 | 644.04 | 686.61 | 787.68 |
| ave. price | 753.82 | 749.42 | 749.50 | 751.68 | 746.33 | 747.43 | 755.42 |
| **Market Share** | | | | | | | |
| FFS | 78.05% | 94.74% | 94.89% | 96.40% | 76.52% | 76.65% | 78.77% |
| 1st MA plan | 2.58% | 0.37% | 0.36% | 0.25% | 2.76% | 2.23% | 1.90% |
| 2nd MA plan | 2.34% | 0.37% | 0.36% | 0.25% | 2.43% | 2.52% | 1.91% |
| 3rd MA plan | 2.63% | 0.37% | 0.36% | 0.26% | 2.19% | 2.19% | 1.76% |
| **Out-of-pocket payment per month** | | | | | | | |
| 0 | 89.12% | 1.52% | 11.50% | 69.49% | 17.34% | 28.89% | 68.44% |
| (0,40] | 8.07% | 12.67% | 18.41% | 16.53% | 16.03% | 21.21% | 16.69% |
| (40, 70] | 2.03% | 12.11% | 14.41% | 8.53% | 14.15% | 14.47% | 8.14% |
| (70,125] | 0.68% | 22.93% | 19.85% | 5.00% | 22.52% | 17.56% | 5.94% |
| (125,200] | 0.07% | 22.58% | 16.30% | 0.45% | 14.93% | 10.70% | 0.74% |
| $> 200$ | 0.03% | 28.18% | 19.53% | 0.01% | 15.04% | 7.17% | 0.05% |
| ave. payment | 3.37 | 146.70 | 107.95 | 12.54 | 102.29 | 66.83 | 13.76 |
| **Government payment ($) per month** | | | | | | | |
| payment | 815.87 | 599.96 | 641.82 | 775.75 | 641.80 | 683.50 | 776.12 |
| % of Exogenous | | 73.54% | 78.67% | 95.08% | 78.66% | 83.78% | 95.13% |

Table B.7: Estimation Result via Estimation (Plan I and inertia cost $S = 4$)

| | | Estimation Model | | | Calibration Model | | |
|---|---|---|---|---|---|---|---|
| | Exogenous | Lowest | Second | Exogenous ($C - 50$) | Lowest | Second | Exogenous ($C - 50$) |
| **Prices** | | | | | | | |
| FFS | 752.96 | 754.74 | 754.70 | 750.53 | 757.09 | 757.64 | 753.13 |
| Capitation | 837.68 | 599.96 | 642.65 | 787.68 | 641.80 | 685.02 | 787.68 |
| ave. price | 753.82 | 754.81 | 754.83 | 752.93 | 746.61 | 749.79 | 760.50 |
| **Market Share** | | | | | | | |
| FFS | 78.05% | 99.49% | 99.21% | 95.70% | 91.80% | 89.52% | 82.48% |
| 1st MA plan | 2.58% | 0.05% | 0.08% | 0.32% | 1.35% | 1.35% | 1.60% |
| 2nd MA plan | 2.34% | 0.05% | 0.08% | 0.32% | 1.04% | 1.31% | 1.57% |
| 3rd MA plan | 2.63% | 0.05% | 0.07% | 0.31% | 1.11% | 1.46% | 1.40% |
| **Out-of-pocket payment per month** | | | | | | | |
| 0 | 97.38% | 92.80% | 91.77% | 95.89% | 99.49% | 99.21% | 99.12% |
| (0,40] | 1.82% | 1.61% | 2.83% | 2.31% | 0.00% | 0.03% | 0.49% |
| (40, 70] | 0.45% | 1.52% | 1.85% | 0.58% | 0.01% | 0.12% | 0.39% |
| (70,125] | 0.26% | 2.71% | 2.66% | 0.95% | 0.15% | 0.18% | 0.00% |
| (125,200] | 0.06% | 1.20% | 0.82% | 0.23% | 0.09% | 0.24% | 0.00% |
| $> 200$ | 0.03% | 0.17% | 0.07% | 0.03% | 0.25% | 0.21% | 0.00% |
| ave. payment | 0.92 | 6.10 | 5.47 | 1.99 | 0.94 | 1.11 | 0.29 |
| **Government payment ($) per month** | | | | | | | |
| payment | 815.87 | 753.88 | 753.72 | 755.73 | 748.24 | 749.49 | 760.68 |
| % of Exogenous | | 92.40% | 92.38% | 92.63% | 91.71% | 91.86% | 93.24% |

# Appendix C

# Appendix for Chapter 3

## C.1 Selected Proofs

**Proof of Proposition 7.** Taking derivative with respect to $c_i$, one has:

$$\frac{\partial \pi_i^{NDFA}(b_i, c_i, \beta)}{\partial c_i} \;=\; -\bar{F}^{N-1}(\beta^{-1}(b_i))\mathbb{P}\left\{b_i \leq c_j + X + Z_j\right\}$$

Applying the envelope theorem [Milgrom, 2004], one has that:

$$\pi_i(\beta(c_i), c_i, \beta) \;=\; \pi_i(\beta(\bar{c}), \bar{c}, \beta) - \int_{c_i}^{\bar{c}} \frac{\partial \pi_i(\beta(c), c, \beta)}{\partial c} dc.$$

The rest of the proof follows by taking $\pi_i(\beta(\bar{c}), \bar{c}, \beta) = 0$ and equating the above with

$$\pi_i(\beta(c_i), c_i, \beta) = \bar{F}^{N-1}(c_i)\mathbb{E}\left[\mathbb{I}\{\beta(c_i) \leq c_j + X + Z_j\} \cdot (\beta(c_i) - c_i - X)\right].$$ □

**Proof of Proposition 8.** Taking derivatives of bidder $i$'s profit with respect to $c_i$, and using

$$\frac{\partial \mathbb{E}\left[\min\{b_i, X + c_i + Z_i\} - c_i - X\right]}{\partial c_i} \;=\; \mathbb{E}\left[-\mathbb{I}\{b_i \leq X + c_i + Z_i\}\right],$$

one obtains the partial derivative specified in the proposition. Applying the envelope theorem, one has that:

$$\pi_i(\beta(c_i), c_i, \beta) \;=\; \pi_i(\beta(\bar{c}), \bar{c}, \beta) - \int_{c_i}^{\bar{c}} \frac{\partial \pi_i(\beta(c), c, \beta)}{\partial c} dc.$$

The bidder's equilibrium profit can be simplified as:

$$\begin{aligned}
\pi_i(\beta(c_i), c_i, \beta) &= \mathbb{E}\Bigg[\mathbb{I}\{c_i \le c_{(1:N),-i}\}\left(\sum_{j=1}^{M} q_{j,M}\mathbb{I}\{\beta(c_i) \le X + c_j + Z_j\}\right)\cdot(\beta(c_i) - c_i - X) \\
&\quad + q_{i,M} Z_i \mathbb{I}\{c_i \le c_{(1:N),-i}, \beta(c_i) > X + c_i + Z_i\} \\
&\quad + q_{i,M} Z_i \mathbb{I}\{c_i > c_{(1:N),-i}, \beta(c_{(1:N),-i}) > X + c_i + Z_i\}\Bigg] \\
&= \mathbb{E}\Bigg[\mathbb{I}\{c_i \le c_{(1:N),-i}\}\left(\sum_{j=1}^{M} q_{j,M}\mathbb{I}\{\beta(c_i) \le X + c_j + Z_j\}\right)\cdot(\beta(c_i) - c_i - X) \\
&\quad + q_{i,M} Z_i \mathbb{I}\{\beta(\min\{c_i, c_{(1:N),-i}\}) > X + c_i + Z_i\}\Bigg].
\end{aligned}$$

Also, the profit associated with the highest cost $\bar{c}$ is strictly positive because of the potential value available for suppliers at the spot market: $\pi_i(\beta(\bar{c}), \bar{c}, \beta) = q_{i,M}\mathbb{E}\left[Z_i \mathbb{I}\{\beta(c_{(1:N),-i}) > X + \bar{c} + Z_i\}\right]$. All together, one obtains the integral equations in the proposition. This concludes the proof. □

**Proof of Theorem 10.** By Proposition 10, we have $\mathbb{E}[P^{NFA}] \ge \sum_{t=1}^{T} P_t^{NFA}$. By standard arguments based on the envelope theorem [Milgrom, 2004], it can be shown that the expected payment under first-price auction is given by $\mathbb{E}[P^{FPA}] = \sum_{t=1}^{T} \mathbb{E}\left[c_{(1:N)} + X_t + F(c_{(1:N)})/f(c_{(1:N)})\right]$. Let $P_t^{FPA} = \mathbb{E}\left[c_{(1:N)} + X_t + F(c_{(1:N)})/f(c_{(1:N)})\right]$. Denote the event of buying from the FA winner under naive FA by $\mathbb{I}^{NFA} := \mathbb{I}\{\beta(c_{(1:N)}) < p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)\}$. By Proposition 10, one has (using the notation $\mu_c = \mathbb{E}[c_i]$):

$$\begin{aligned}
P_t^{NFA} &= \sum_{i=1}^{N} \pi_{i,t}(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}) + \mathbb{E}\left[\widetilde{q}_0(\boldsymbol{c}, X_t)\right] + \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t)\left(v(c_i) + X_t - \widetilde{q}_0(\boldsymbol{c}, X_t)\right)\right] \\
&\overset{(a)}{=} \sum_{i=1}^{N} \mathbb{E}[A_i Z_{i,t}(1 - \mathbb{I}^{NFA})] + \mathbb{E}\left[\widetilde{q}_0(\boldsymbol{c}, X_t)\right] + \mathbb{E}\left[\mathbb{I}^{NFA}\left(v(c_{(1:N)}) + X_t - \widetilde{q}_0(\boldsymbol{c}, X_t)\right)\right] \\
&\overset{(b)}{=} \sum_{i=1}^{N} \mathbb{E}[A_i Z_{i,t}(1 - \mathbb{I}^{NFA})] + \mathbb{E}\left[\sum_{i=1}^{M} A_i(c_i + X_t)(1 - \mathbb{I}^{NFA})\right] + \mathbb{E}\left[\mathbb{I}^{NFA}\left(v(c_{(1:N)}) + X_t\right)\right],
\end{aligned}$$

where (a) follows from $\pi_{i,t}(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}) = \mathbb{E}[A_i Z_{i,t}(1 - \mathbb{I}^{NFA})]$ and $r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t) = \mathbb{I}\{\beta(c_i) < \beta(c_j), \forall j \ne i, \beta(c_i) < p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)\}$, and (b) follows by $\widetilde{q}_0(\boldsymbol{c}, X_t) = \sum_{i=1}^{M} A_i(c_i + X_t)$. Substituting the expression of $P_t^{FPA}$, one has:

$$\begin{aligned}
P_t^{NFA} - P_t^{FPA} &= \sum_{i=1}^{N} \mathbb{E}[A_i Z_{i,t}(1 - \mathbb{I}^{NFA})] + \mathbb{E}\left[\sum_{i=1}^{M} A_i(c_i + X_t)(1 - \mathbb{I}^{NFA})\right] \\
&\quad + \mathbb{E}\left[\mathbb{I}^{NFA}\left(v(c_{(1:N)}) + X_t\right)\right] - \mathbb{E}\left[v(c_{(1:N)}) + X_t\right] \\
&\overset{(a)}{=} \mathbb{E}\left[\left(\sum_{i=1}^{N} A_i Z_{i,t} + \sum_{i=1}^{M} A_i c_i - v(c_{(1:N)})\right)(1 - \mathbb{I}^{NFA})\right], \qquad \text{(C-1)}
\end{aligned}$$

where (a) follows by the fact that $\sum_{i=1}^{M} A_i = 1$.

**1. Concentrated Markets.** Without loss of generality, we assume $M = N$. One has

$$\begin{aligned}\lim_{N\to\infty}\left[P_t^{NFA} - P_t^{FPA}\right] &= \lim_{N\to\infty}\mathbb{E}\left[\left(\sum_{i=1}^{N} A_i Z_{i,t} + \sum_{i=1}^{N} A_i c_i - v(c_{(1:N)})\right)(1-\mathbb{I}^{NFA})\right] \\ &\overset{(a)}{=} \lim_{N\to\infty}\mathbb{E}\left[\left(\sum_{i=1}^{N} A_i (Z_{i,t} + c_i) - \underline{c}\right)(1-\mathbb{I}^{NFA})\right] \geq 0,\end{aligned}$$

where $(a)$ follows from applying bounded convergence theorem to $\lim_{N\to\infty}\mathbb{E}\left[\left(v(c_{(1:N)}) - \underline{c}\right)(1-\mathbb{I}^{NFA})\right] = 0$, because $c_{(1:N)}$, the lowest private cost among the $N$ participants of the auction stage, converges to $\underline{c}$ in probability and $\frac{F(c_{(1:N)})}{f(c_{(1:N)})} \to 0$ in probability when $N \to \infty$. The inequality follows from $\sum_{i=1}^{N} A_i(Z_{i,t} + c_i) - \underline{c} = \sum_{i=1}^{N} A_i(Z_{i,t} + c_i - \underline{c}) \geq 0$ and $1 - \mathbb{I}^{NFA} \geq 0$ for any sample path.

**2. Diffused Markets.** We note that $\lim_{M\to\infty}\sum_{i=1}^{N} q_{i,M} = 0$. By bounded convergence theorem, one has:

$$\lim_{N\to\infty}\lim_{M\to\infty} \quad \mathbb{E}\left[\sum_{i=1}^{N} A_i Z_{i,t}(1-\mathbb{I}^{NFA})\right] = 0, \tag{C-2}$$

where the order of the limits is consistent with the diffused market assumption.

Note that $v(c_{(1:N)}) - c_{(1:N)} = \frac{F(c_{(1:N)})}{f(c_{(1:N)})}$ converges in probability to 0 (recall that $f(\underline{c}) > 0$). Then, by bounded convergence theorem we have:

$$\begin{aligned}&\lim_{N\to\infty}\lim_{M\to\infty}\mathbb{E}\left[\left(\sum_{i=1}^{M} A_i c_i - v(c_{(1:N)})\right)(1-\mathbb{I}^{NFA})\right] \\ =\ &\lim_{N\to\infty}\lim_{M\to\infty}\mathbb{E}\left[\left(\sum_{i=1}^{M} A_i c_i - c_{(1:N)}\right)(1-\mathbb{I}^{NFA})\right] \\ =\ &\lim_{N\to\infty}\mathbb{E}\left[\left(c - c_{(1:N)}\right)\mathbb{I}\left\{\beta_\infty^{NFA}(c_{(1:N)}) \geq c + Z_t + X_t\right\}\right] \\ =\ &\mathbb{E}\left[(c - \underline{c})\,\mathbb{I}\left\{\beta_\infty^{NDFA}(\underline{c}) \geq c + Z_t + X_t\right\}\right]\end{aligned} \tag{C-3}$$

Here, $\beta_M^{NDFA}(\cdot)$ is the bidding strategy under naive FA with $M$ potential suppliers. The second equation follows since $(c, Z_t)$ has the same distribution as $(c_i, Z_{i,t})$, which has the same marginal distribution for all $i$, and it is independent of $c_{(1:N)}$ by the diffused market assumption. The last equation holds by bounded convergence theorem. Extending Proposition 7 to multiple periods:

$$\begin{aligned}
\beta^{NDFA}(c) &= c + \frac{\sum_{t=1}^{T} \mathbb{E}\left[X_t \mathbb{I}\left\{\beta^{NDFA}(c) \le c_j + X_t + Z_{j,t}\right\}\right]}{\sum_{t=1}^{T} \mathbb{P}\left[\beta^{NDFA}(c) \le c_j + X_t + Z_{j,t}\right]} \\
&+ \frac{\sum_{t=1}^{T} \int_{c}^{\bar{c}} \bar{F}^{N-1}(y) \cdot \mathbb{P}\left\{\beta^{NDFA}(y) \le c_j + X_t + Z_{j,t}\right\} dy}{\bar{F}^{N-1}(c) \cdot \sum_{t=1}^{T} \mathbb{P}\left\{\beta^{NDFA}(c) \le c_j + X_t + Z_{j,t}\right\}} \\
&\ge c + \min_t \frac{\mathbb{E}\left[X_t \mathbb{I}\left\{\beta^{NDFA}(c) \le c_j + X_t + Z_{j,t}\right\}\right]}{\mathbb{P}\left[\beta^{NDFA}(c) \le c_j + X_t + Z_{j,t}\right]} \ge c + \mathbb{E}[X_{\hat{t}}], \qquad \text{(C-4)}
\end{aligned}$$

where $\hat{t}$ achieves the minimum. Hence:

$$\begin{aligned}
\lim_{N\to\infty} \mathbb{E}\left[P^{NDFA} - P^{FPA}\right] &\ge \lim_{N\to\infty} \sum_{t=1}^{T} \left(P_t^{NFA} - P_t^{FPA}\right) \\
&= \sum_{t=1}^{T} \mathbb{E}\left[(c - \underline{c})\, \mathbb{I}\left\{\beta_{\infty}^{NDFA}(\underline{c}) \ge c + Z_t + X_t\right\}\right] \\
&\ge \mathbb{E}\left[(c - \underline{c})\, \mathbb{I}\left\{\beta_{\infty}^{NDFA}(\underline{c}) \ge c + Z_{\hat{t}} + X_{\hat{t}}\right\}\right] \\
&\ge \mathbb{E}\left[(c - \underline{c})\, \mathbb{I}\left\{\underline{c} + \mathbb{E}[X_{\hat{t}}] \ge c + Z_{\hat{t}} + X_{\hat{t}}\right\}\right] > 0, \qquad \text{(C-5)}
\end{aligned}$$

where the second expression follows by (C-1), (C-2) and (C-3), the third because $c - \underline{c} \ge 0$, the fourth by (C-4), and the last because $\underline{z} + \underline{x} < \mathbb{E}[X_t]$. This concludes the proof. □

**Proof of Proposition 11.** The proof follows ideas similar to the ones described in the proof of Proposition 10. Consider the expected total profits from period 1 to period $T$ for bidder $i$ under bid $b_i$, cost $c_i$, and when competitors use equilibrium strategy $\boldsymbol{\beta}_{-i}$:

$$\begin{aligned}
\pi_i(b_i, c_i, \boldsymbol{\beta}_{-i}) &= \sum_{t=1}^{T} \pi_{i,t}(b_i, c_i, \boldsymbol{\beta}_{-i}), \text{ where} \\
\pi_{i,t}(b_i, c_i, \boldsymbol{\beta}_{-i}) &= \mathbb{E}_{-i}\Big[ (b_i - c_i - X_t)\, r_{i,t}(b_i, \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), p_i(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)) \\
&\quad + [p_i(\boldsymbol{c}, X_t, Z_{i,t}) - c_i - X_t]\, \mathbb{I}\{b_i < \beta(c_j), \forall j \ne i\}\, \mathbb{I}\{b_i \ge p_i(\boldsymbol{c}, X_t, Z_{i,t})\} \Big],
\end{aligned}$$

where the first term is the profit in the event that bidder $i$ is the FA winner and his bid is below his own spot market price, and the second term is the profit when he is the FA winner but loses to his own spot market price. Substituting $p_i(\boldsymbol{c}, X_t, \boldsymbol{Z}_t) = c_i + X_t + Z_{i,t}$ into the above, one has:

$$\begin{aligned}
\pi_{i,t}(b_i, c_i, \boldsymbol{\beta}_{-i}) &= \mathbb{E}_{-i}\bigg[ (b_i - c_i - X_t)\,\mathbb{I}\{b_i < \beta(c_j), \forall j \neq i\}\,\mathbb{I}\{b_i < c_i + X_t + Z_{i,t}\} \\
&\quad + Z_{i,t}\mathbb{I}\{b_i < \beta(c_j), \forall j \neq i\}\,\mathbb{I}\{b_i \geq c_i + X_t + Z_{i,t}\}\bigg] \\
&= \mathbb{E}_{-i}\left[\mathbb{I}\{b_i < \beta(c_j), \forall j \neq i\} \cdot \int \left( \int_{b_i - c_i - z} (b_i - c_i - x)\, f_{X_t}(x)dx \right) f_{Z_{i,t}}(z)dz \right] \\
&\quad + \mathbb{E}_{-i}\left[ \{b_i < \beta(c_j), \forall j \neq i\} \cdot \int \left( \int^{b_i - c_i - z} z f_{X_t}(x)dx \right) f_{Z_{i,t}}(z)dz \right].
\end{aligned}$$

Therefore, taking derivative with respect to $c_i$, one has:

$$\begin{aligned}
\frac{\partial \pi_{i,t}(b_i, c_i, \boldsymbol{\beta}_{-i})}{\partial c_i}\Big|_{b_i=\beta(c_i)} &\overset{(a)}{=} \mathbb{E}_{-i}\bigg[\mathbb{I}\{c_i < c_j, \forall j \neq i\} \cdot \left(-\bar{F}_{X_t}(b_i - c_i - Z_{i,t}) + Z_{i,t} f_{X_t}(b_i - c_i - Z_{i,t})\right) \\
&\quad - \mathbb{I}\{c_i < c_j, \forall j \neq i\}\, Z_{i,t} f_{X_t}(b_i - c_i - Z_{i,t})\bigg] \\
&\overset{(b)}{=} \mathbb{E}_{-i}\left[-\mathbb{I}\{c_i < c_j, \forall j \neq i\}\, \bar{F}_{X_t}(b_i - c_i - Z_{i,t})\right] \\
&= \mathbb{E}_{-i}\left[-\mathbb{I}\{c_i < c_j, \forall j \neq i\}\,\mathbb{I}\{b_i < c_i + X_t + Z_{i,t}\}\right] \\
&\overset{(c)}{=} \mathbb{E}_{-i}\left[-r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_i(\boldsymbol{c}, X_t, \boldsymbol{Z}_t))\right],
\end{aligned}$$

where: $(a)$ follows from applying Leibniz rule; $(b)$ follows by simplifying terms; and $(c)$ follows from the definition of the equilibrium allocation in a monitored FA, in which the FA winner is the lowest cost supplier given that equilibrium bids are strictly increasing. Let $p_{i,t} = p_i(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)$, applying the envelop theorem yields:

$$\begin{aligned}
\pi_i(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) - \pi_i(\beta(c_i), c_i, \boldsymbol{\beta}_{-i}) &= \int_{c_i}^{\bar{c}} \frac{\partial \pi_i(b_i, s, \boldsymbol{\beta})}{\partial s}\Big|_{b_i=\beta(s)} ds \\
&= \sum_{t=1}^{T} \int_{c_i}^{\bar{c}} \mathbb{E}_{-i}\left[-r_{i,t}(\beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), p_{i,t})\right] ds,
\end{aligned}$$

where we omit the arguments of $p(\cdot)$ to simplify notation. Note that

$$\pi_i(\beta(c_i), c_i, \boldsymbol{\beta}_{-i}) = \sum_{t=1}^{T} \mathbb{E}_{-i}\Big[ m_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t}) - (c_i + X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t}) + [p_{i,t} - c_i - X_t]\,\mathbb{I}\{c_i < c_j, \forall j \neq i\}\,\mathbb{I}\{\beta(c_i) \geq p_{i,t}\}$$

Combining the above two equations, and using the fact that $\pi_{i,t}(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) = 0$, one has

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}_{-i}\left[m_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t})\right] &= \sum_{t=1}^{T} \bigg\{ \mathbb{E}_{-i}\Big[ (c_i + X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t}) - [p_{i,t} - c_i - X_t]\,\mathbb{I}\{c_i < c_j, \forall j \neq i\}\,\mathbb{I}\{\beta(c_i) \geq p_{i,t}\}\Big] \\
&\quad + \int_{c_i}^{\bar{c}} \mathbb{E}_{-i}\left[ r_{i,t}(\beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), p_{i,t})\right] ds \bigg\}.
\end{aligned}$$

If the auctioneer does not buy from one of the FA bidders in the FA, she buys from the spot market. Therefore, the expected total buying price for the monitored FA from period 1 to period $T$ is:

$$\mathbb{E}[P^{MFA}] = \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N} m_{i,t}(\beta(\boldsymbol{c}), p_{i,t}) + \sum_{t=1}^{T}\sum_{i=1}^{N} \mathbb{I}\{c_i < c_j, \forall j \neq i\}\,\mathbb{I}\{\beta(c_i) \geq p_{i,t}\}\, p_{i,t}\right] = \sum_{t=1}^{T} P_t^{MFA},$$

where

$$\begin{aligned}
P_t^{MFA} &= \sum_{i=1}^{N}\Bigg\{\mathbb{E}\Bigg[(c_i + X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t}) - [p_{i,t} - c_i - X_t]\,\mathbb{I}\{c_i < c_j, \forall j \neq i\}\,\mathbb{I}\{\beta(c_i) \geq p_{i,t}\}\Bigg] \\
&\quad + \int_{c_i}^{\bar{c}} \mathbb{E}\left[r_{i,t}(\beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), p_{i,t})\right] ds \;+\; \mathbb{E}\left[\mathbb{I}\{c_i < c_j, \forall j \neq i\}\,\mathbb{I}\{\beta(c_i) \geq p_{i,t}\}\, p_{i,t}\right]\Bigg\} \\
&= \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t})\,(c_i + X_t)\right] + \mathbb{E}\left[\sum_{i=1}^{N}(c_i + X_t)\mathbb{I}\{c_i < c_j, \forall j \neq i\}\,\mathbb{I}\{\beta(c_i) \geq p_{i,t}\}\right] \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{N}\int_{c_i}^{\bar{c}} r_{i,t}(\beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), p_{i,t}) ds\right] \\
&\overset{(a)}{=} \mathbb{E}\left[\sum_{i=1}^{N}(c_i + X_t)\mathbb{I}\{c_i < c_j, \forall j \neq i\}\left(1 - \sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t})\right)\right] \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t})\left(c_i + \frac{F(c_i)}{f(c_i)} + X_t\right)\right],
\end{aligned}$$

where the last equation is established by changing the order of integration and since

$$(1 - \sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t}))\mathbb{I}\{c_i < c_j, \forall j \neq i\} = \mathbb{I}\{\beta(c_i) \geq p_{i,t}\}\,\mathbb{I}\{c_i < c_j, \forall j \neq i\}.$$

Since $\sum_{i=1}^{N}(c_i + X_t)\mathbb{I}\{c_i < c_j, \forall j \neq i\} = c_{(1:N)} + X_t = q_0(\boldsymbol{c}, X_t)$, one has

$$P_t^{MFA} \;=\; \mathbb{E}\left[q_0(c, X_t)\right] + \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_{i,t})\,(v(c_i) + X_t - q_0(\boldsymbol{c}, X_t))\right].$$

This concludes the proof. □

**Proof of Theorem 11.** Using the envelope theorem we obtain $E\left[P^{FPA}\right] = \sum_{t=1}^{T} P_t^{FPA}$, where $P_t^{FPA} = \mathbb{E}\left[c_{(1:N)} + X_t + F(c_{(1:N)})/f(c_{(1:N)})\right]$. By Proposition 11, $E\left[P^{MFA}\right] = \sum_{t=1}^{T} P_t^{MFA}$, and thus $\mathbb{E}\left[P^{MFA} - P^{FPA}\right] = \sum_{t=1}^{T}\left(P_t^{MFA} - P_t^{FPA}\right)$. One has:

$$\begin{aligned}
P_t^{MFA} - P_t^{FPA} &= \mathbb{E}\left[q_0(\boldsymbol{c}, X_t)\right] \;+\; \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t))\,(v(c_i) + X_t - q_0(\boldsymbol{c}, X_t))\right] \\
&\quad - \mathbb{E}\left[c_{(1:N)} + X_t + F(c_{(1:N)})/f(c_{(1:N)})\right] \\
&\overset{(a)}{=} \mathbb{E}\left[(\mathbb{I}_t^{MFA} - 1)F(c_{(1:N)})/f(c_{(1:N)})\right] \;\leq\; 0,
\end{aligned}$$

where $\mathbb{I}_t^{MFA} = \mathbb{I}\{\beta^{MFA}(c_{(1:N)}) < c_{(1:N)} + X_t + Z_{(1:N),t}\}$ is the event that FA winner wins over spot market, and where (a) follows from $q_0(\boldsymbol{c}, x) = c_{(1:N)} + x$, $v(c_i) = c_i + F(c_i)/f(c_i)$, and

$r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)) = \mathbb{I}\left\{c_i < c_j, \forall j \neq i, \beta(c_i) < c_i + X_t + Z_{i,t}\right\}$. We note that $c_{(1:N)}$, the lowest private cost among the $N$ participants of the auction stage, converges to $\underline{c}$ in probability as $N \to \infty$. In addition, $\frac{F(c_{(1:N)})}{f(c_{(1:N)})}$ converges in probability to 0 (recall that $f(\underline{c}) > 0$). Since $|\mathbb{I}^{MFA}| \leq 1$, by bounded convergence theorem, we have:

$$\mathbb{E}\left[(\mathbb{I}_t^{MFA} - 1)F(c_{(1:N)})/f(c_{(1:N)})\right] \longrightarrow 0 \quad \text{as} \quad N \longrightarrow \infty.$$

The result follows. □

**Proof of Theorem 12.** From Proposition 11 and Proposition 12, one has $\mathbb{E}[P^{FLE} - P^{MFA}] = \sum_{t=1}^{T}(P_t^{FLE} - P_t^{MFA})$, where

$$\begin{aligned} P_t^{FLE} - P_t^{MFA} &= \mathbb{E}[Z_t] + \mathbb{E}\left[\mathbb{I}_t^{FLE}(v(c_{(1:N)}) - Z_t)\right] - \mathbb{E}\left[c_{(1:N)}\right] - \mathbb{E}\left[\mathbb{I}_t^{MFA}\left(v(c_{(1:N)}) - c_{(1:N)}\right)\right] \\ &= \mathbb{E}\left[(1 - \mathbb{I}_t^{FLE})(Z_t - c_{(1:N)})\right] + \mathbb{E}\left[(\mathbb{I}_t^{FLE} - \mathbb{I}_t^{MFA})\left(v(c_{(1:N)}) - c_{(1:N)}\right)\right], \end{aligned}$$

where $\mathbb{I}_t^{FLE} = \mathbb{I}\left\{c_{(1:N)} \leq Z_t\right\} + \mathbb{I}\left\{\beta^{FLE}(c_{(1:N)}) \leq Z_t + X_t, c_{(1:N)} > Z_t\right\}$ and $\mathbb{I}_t^{MFA} = \mathbb{I}\left\{\beta^{MFA}(c_{(1:N)}) \leq c_{(1:N)} + X_t + Z_{(1:N),t}\right\}$. Since $c_{(1:N)}$ converges to $\underline{c}$ in probability as $N \to \infty$ and $\underline{c} \leq \underline{z}_t$ (recall $Z_t = \sum_{j=1}^{\infty} A_j(c_j + Z_{j,t})$), $1 - \mathbb{I}_t^{FLE}$ converges in probability to zero, thus, by bounded convergence theorem, $\mathbb{E}\left[(1 - \mathbb{I}_t^{FLE})(Z_t - c_{(1:N)})\right]$ converges to zero. In addition, $v(c_{(1:N)}) - c_{(1:N)} = \frac{F(c_{(1:N)})}{f(c_{(1:N)})}$ converges in probability to 0 (recall that $f(\underline{c}) > 0$), and $|\mathbb{I}_t^{FLE} - \mathbb{I}_t^{MFA}| \leq 2$, by bounded convergence theorem, $\mathbb{E}\left[(\mathbb{I}_t^{FLE} - \mathbb{I}_t^{MFA})\left(v(c_{(1:N)}) - c_{(1:N)}\right)\right]$ converges to zero.

Next, we show that $\mathbb{E}[P^{MFA} - P^{FPA}] = \sum_{t=1}^{T}(P_t^{MFA} - P_t^{FPA}) \to 0$ as $N \to \infty$. From the proof of Theorem 11, one has

$$P_t^{MFA} - P_t^{FPA} = \mathbb{E}\left[(\mathbb{I}_t^{MFA} - 1)F(c_{(1:N)})/f(c_{(1:N)})\right].$$

Again, as $N \to \infty$, $P_t^{MFA} - P_t^{FPA} \to 0$ by the bounded convergence theorem. Finally, we show that $\mathbb{E}[P^{FLE} - P^{OPT}] \to 0$ as $N \to \infty$. According to Appendix C.6, one has

$$\mathbb{E}[P^{OPT}] = \sum_{t=1}^{T} P_t^{OPT}, \qquad \text{where,} \qquad P_t^{OPT} = \mathbb{E}\left[Z_t + X_t + \min\{0, v(c_{(1:N)}) - Z_t\}\right].$$

Together with Proposition 12, one has

$$\begin{aligned} P_t^{FLE} - P_t^{OPT} &= \mathbb{E}\left[\mathbb{I}_t^{FLE}(v(c_{(1:N)}) - Z_t)\right] - \mathbb{E}\left[\min\{0, v(c_{(1:N)}) - Z_t\}\right] \\ &= \mathbb{E}\left[(\mathbb{I}_t^{FLE} - \mathbb{I}_t^{OPT})(v(c_{(1:N)}) - Z_t)\right], \end{aligned}$$

where $\mathbb{I}_t^{OPT} = \mathbb{I}\{v(c_{(1:N)}) \leq Z_t\}$. As $N \to \infty$, since $v(c_{(1:N)})$ convergence to $\underline{c}$ in probability, and $\underline{c} \leq \underline{z}_t$, $\mathbb{I}_t^{OPT} \to 1$ and $\mathbb{I}_t^{FLE} \to 1$ in probability. Thus, $P_t^{FLE} - P_t^{OPT} \to 0$, by the bounded convergence theorem. This completes the proof. □

## C.2 Ordinary Differential Equations for Numerical Experiments

The following result establishes the ordinary differential equations (ODEs) and the corresponding boundary conditions that characterize symmetric BNE strategies.

**Proposition C1 (ODEs associated with FA mechanisms)** *Let $z_0 = c_0 + \Delta$. Then,*

*(1) A symmetric and differentiable BNE strategy under naive FA in diffused market satisfies the following ODE:*

$$\frac{d\beta^{NDFA}(c)}{dc} = \frac{(N-1)f(c)}{\bar{F}(c)} \cdot \frac{\mathbb{E}_X\left[\left(\beta^{NDFA}(c) - c - X\right) \cdot \mathbb{I}\left\{\beta^{NDFA}(c) \leq z_0 + X\right\}\right]}{\left[\bar{F}_X\left(\beta^{NDFA}(c) - z_0\right) - (z_0 - c) f_X(\beta^{NDFA}(c) - z_0)\right]},$$

*for any $c \in [\underline{c}, z_0]$, with the boundary conditions $\beta^{NDFA}(z_0) = z_0 + \overline{x}$, and $\frac{d\beta^{NDFA}}{dc}(z_0) = 0$.*

*(2) A symmetric and differentiable BNE strategy under monitored FA satisfies the following ODE:*

$$\frac{d\beta^{MFA}(c)}{dc} = \frac{(N-1)f(c)}{\bar{F}(c)} \cdot \frac{\mathbb{E}_X\left[\min\{\beta^{MFA}(c) - c - X, \Delta\}\right]}{\mathbb{P}[\beta^{MFA}(c) \leq c + \Delta + X]},$$

*for any $c \in [\underline{c}, \bar{c}]$, with the boundary condition $\beta^{MFA}(\bar{c}) = \bar{c} + \Delta + K$, where $K = \mathbb{E}[X] - \Delta$ if $\mathbb{E}[X] \leq \Delta$, and otherwise $K \in [0, \bar{x}]$ is the unique solution to the equation:*

$$(K + \Delta)\bar{F}_X(K) - \int_K^{\bar{x}} x f_X(x) dx + \Delta \cdot F_X(K) = 0.$$

*(3) A symmetric and differentiable BNE strategy under flexible FA satisfies the following ODE:*

$$\frac{d\beta^{FLE}(c)}{dc} = \frac{(N-1)f(c)}{\bar{F}(c)} \cdot \left.\frac{\mathbb{E}_X\left[(b - c - X) \cdot \mathbb{I}\left\{b \leq z_0 + X\right\} + (z_0 - c) \cdot \mathbb{I}\left\{b > z_0 + X\right\}\right]}{\mathbb{P}[b \leq z_0 + X]}\right|_{b=\beta^{FLE}(c)},$$

*for any $c \in [\underline{c}, z_0]$, with the boundary conditions $\beta^{FLE}(z_0) = z_0 + \overline{x}$, and $\frac{d\beta^{FLE}}{dc}(z_0) = 0$.*

*(4) Assume $T = 2$. Then, a symmetric and differentiable BNE strategy under restricted-flexible FA satisfies the ODE:*

$$\frac{d\beta^{FLR}(c)}{dc} = \frac{(N-1)f(c) \cdot V_1(\beta(c), c)}{\overline{F}(c) \cdot \left.\frac{\partial V_1(b,c)}{\partial b}\right|_{b=\beta(c)}},$$

*where*

$$V_1(b_0, c) = \int_{B(b_0,z_0,c)}^{b_0 - z_0} \left[(z_0 - c) + V_2(x + z_0, c) - V_2(b_0, c)\right] F_X(x) dx$$
$$+ V_2(b_0, c) + \mathbb{E}\left[(b_0 - c - X_1)\,\mathbb{I}\{b_0 \leq X_1 + z_0\}\right],$$

*and*

$$\begin{aligned} V_2(b_1,c) &= \mathbb{E}\left[(b_1-c-X_2)\mathbb{I}\{b_1 \le X_2+z_0\} + (z_0-c)\mathbb{I}\{b_1 > X_2+z_0\}\right], \\ B(b_0,z_0,c) &= \inf\{x\in[0,\bar{x}] : V_2(b_0,c)-(z_0-c)-V_2(x+z_0,c) \le 0.\}. \end{aligned}$$

*for any* $c \in [\underline{c}, z_0]$*, with the boundary conditions* $\beta^{FLR}(z_0) = z_0+\overline{x}$*, and* $\frac{d\beta^{FLR}}{dc}(z_0) = 0$.

We note that if $X$ is uniformly distributed over the interval $[0,\bar{x}]$, the boundary condition under monitored FA takes the following closed form expression: $\beta^{MFA}(\bar{c}) = \bar{c}+\Delta+\bar{x}-\sqrt{2\Delta\bar{x}}$. We further note that while the FA winner under the naive FA, flexible FA, or restricted-flexible FA is competing against an outside market with price $z_0+X$, and as a result bidders with private cost higher than $z_0$ never win, under the monitored FA the FA winner is competing against his own spot market price, and therefore the whole interval $[\underline{c},\overline{c}]$ should be considered.

Similarly to asymmetric first-price auctions, our ODEs are not well-behaved at the boundary, because at the right-hand-side of these we obtain $\frac{0}{0}$. To avoid the singularity at the boundary, we make the approximation $\beta(z_0-\epsilon) = \beta(z_0)-\epsilon$ for a small value $\epsilon = 10^{-5}$, and a first order approximation yields $\beta(z_0-\epsilon) = \beta(z_0)$, when $\beta'(z_0) = 0$. Since we are looking for BNE in strictly increasing strategies, and to avoid a flat curve at the boundary, we subtract $\epsilon$ from $\beta(z_0)$ above.

## C.3 Additional Proofs

**Proof of Proposition 9.** Under monitored FA, bidder's profit can be written as

$$\begin{aligned} \pi_i(b_i,c_i,\beta) &= \mathbb{E}\left[\mathbb{I}\{b_i \le \beta(c_{(1:N),-i})\}\mathbb{I}\{b_i \le c_i+X+Z_i\}(b_i-c_i-X)\right] \\ &\quad + \mathbb{E}\left[\mathbb{I}\{b_i \le \beta(c_{(1:N),-i})\}Z_i\mathbb{I}\{b_i > c_i+X+Z_i\}\right] \\ &= \bar{F}^{N-1}(\beta^{-1}(b_i))\int\int_{b_i-c_i-z}(b_i-c_i-x)\,f_X(x)f_Z(z)dxdz \\ &\quad + \bar{F}^{N-1}(\beta^{-1}(b_i))\int\int^{b_i-c_i-z}(c_i+x+z)-c_i-x)\,f_X(x)f_Z(z)dxdz. \end{aligned}$$

Taking derivative with respect to $c_i$, one has

$$\left.\frac{\partial \pi_i(b_i,c_i,\beta)}{\partial c_i}\right|_{b_i=\beta(c_i)} = \bar{F}^{N-1}(c_i)\cdot\mathbb{E}\left[-\mathbb{I}\{\beta(c_i) \le c_i+X+Z_i\}\right]$$

Applying the envelop theorem, one has:

$$\begin{aligned}\pi_i(\beta(c_i), c_i, \beta) &= \pi_i(\beta(\bar{c}), \bar{c}, \beta) - \int_{c_i}^{\bar{c}} \frac{\partial \pi_i(\beta(c), c, \beta)}{\partial c} dc \\ &= \pi_i(\beta(\bar{c}), \bar{c}, \beta) \ + \int_{c_i}^{\bar{c}} \bar{F}^{N-1}(c) \cdot \mathbb{E}\left[\mathbb{I}\{\beta(c) \leq c + X + Z_i\}\right] dc.\end{aligned}$$

Bidder's equilibrium profit is given by:

$$\begin{aligned}\pi_i(\beta(c_i), c_i, \beta) &= \bar{F}^{N-1}(c_i)\mathbb{E}\left[\mathbb{I}\{\beta(c_i) \leq c_i + X + Z_i\}(\beta(c_i) - c_i - X)\right] \\ &\quad + \bar{F}^{N-1}(c_i)\mathbb{E}\left[Z_i\mathbb{I}\{\beta(c_i) > c_i + X + Z_i\}\right].\end{aligned}$$

Taking $\pi_i(\beta(\bar{c}), \bar{c}, \beta) = 0$, one obtains:

$$\begin{aligned}\beta^{\mathrm{MFA}}(c_i) &= c_i + \frac{\mathbb{E}\left[X\mathbb{I}\{\beta(c_i) \leq c_i + X + Z_i\}\right]}{\mathbb{E}\left[\mathbb{I}\{\beta(c_i) \leq c_i + X + Z_i\}\right]} \ + \ \frac{\int_{c_i}^{\bar{c}} \bar{F}^{N-1}(c) \cdot \mathbb{E}\left[\mathbb{I}\{\beta(c) \leq c + X + Z_i\}\right] dc}{\bar{F}^{N-1}(c_i)\mathbb{E}\left[\mathbb{I}\{\beta(c_i) \leq c_i + X + Z_i\}\right]} \\ &\quad - \frac{\mathbb{E}\left[Z_i\mathbb{I}\{\beta(c_i) > c_i + X + Z_i\}\right]}{\mathbb{E}\left[\mathbb{I}\{\beta(c_i) \leq c_i + X + Z_i\}\right]}.\end{aligned}$$

This concludes the proof. □

**Proof of Proposition 10.** Note that the total profit for bidder $i$ from periods 1 to period $T$ equals to the sum of the profits across all periods, and same is true for buyer's expected buying prices:

$$\pi_i = \sum_{t=1}^{T} \pi_{i,t} \text{ and } \mathbb{E}\left[P^{NFA}\right] = \sum_{t=1}^{T} P_t^{NFA}.$$

Consider the expected profits for bidder $i$ at period $t$ when he bids $b_i$, his cost is $c_i$, and his competitors use equilibrium strategy $\boldsymbol{\beta}_{-i}$:

$$\begin{aligned}\pi_{i,t}(b_i, c_i, \boldsymbol{\beta}_{-i}) &= \mathbb{E}_{-i}\Bigg[(b_i - c_i - X_t)\, r_{i,t}(b_i, \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)) \\ &\quad + \left[p_i(\boldsymbol{c}, X_t, Z_{i,t}) - c_i - X_t\right] A_i \left(1 - \sum_{j=1}^{N} r_{j,t}(b_i, \boldsymbol{\beta}_{-i}, p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t))\right)\Bigg], \\ &= \sum_{j=1}^{M} q_{j,M}\mathbb{E}_{-i}\left[(b_i - c_i - X_t)\, \mathbb{I}\left\{b_i < \beta_t(c_t), \forall t \neq i, b_i < c_j + X_t + Z_{j,t}\right\}\right] \\ &\quad + q_{i,M}\mathbb{E}_{-i}\left[Z_{i,t}\mathbb{I}\left\{\min\{b_i, \beta_t(c_t)\}_{t\neq i} \geq c_i + X_t + Z_{i,t}\right\}\right],\end{aligned}$$

where the second equation follows from $p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t) = \sum_{j=1}^{M} A_j(c_j + X_t + Z_{j,t})$, and $(1 - \sum_{j=1}^{N} r_{j,t}(\boldsymbol{b}, p_t)) = \mathbb{I}\left\{\min\{b_j\}_{j=1}^{N} \geq p_t\right\}$. Taking derivative with respect to $c_i$, one has:

$$
\begin{aligned}
\frac{\partial \pi_{i,t}(b_i, c_i, \boldsymbol{\beta})}{\partial c_i}|_{b_i=\beta(c_i)} &= \sum_{j\neq i} q_{j,M}\mathbb{E}_{-i}\left[-\mathbb{I}\left\{c_i < c_\ell, \forall \ell \neq i, \beta(c_i) < c_j + X_t + Z_{j,t}\right\}\right] \\
&\quad + q_{i,M}\mathbb{E}_{-i}\left[-\mathbb{I}\left\{c_i < c_\ell, \forall \ell \neq i, \beta(c_i) < c_i + X_t + Z_{i,t}\right\}\right] \\
&\quad + q_{i,M}\mathbb{E}_{-i}\left[-Z_{i,t} f_{X_t}\left(\beta(c_{(1)}) - c_i - Z_{i,t}\right)\mathbb{I}\left\{c_i > c_{(1)}\right\}\right] \\
&= \mathbb{E}_{-i}\left[-\mathbb{I}\left\{c_i < c_\ell, \forall \ell \neq i, \beta(c_i) < p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)\right\}\right] \\
&\quad - q_{i,M}\mathbb{E}_{-i}\left[Z_{i,t} f_{X_t}\left(\beta(c_{(1)}) - c_i - Z_{i,t}\right)\mathbb{I}\left\{c_i > c_{(1)}\right\}\right] \\
&\overset{(a)}{=} \mathbb{E}_{-i}\left[-r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t))\right] \;-\; q_{i,M}\mathbb{E}_{-i}\left[Z_{i,t} f_{X_t}\left(\beta(c_{(1)}) - c_i - Z_{i,t}\right)\mathbb{I}\left\{c_i > c_{(1)}\right\}\right] \\
&\overset{(b)}{\leq} \mathbb{E}_{-i}\left[-r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t))\right],
\end{aligned}
$$

where $(a)$ follows from the definition of the allocation in the monitored FA ($r_{i,t}(\boldsymbol{b}, p) = \mathbb{I}\left\{b_i < b_j, \forall j \neq i, b_i < p_t\right\}$), which in equilibrium implies the FA winner is the lowest cost suppliers given that equilibrium strategies are strictly increasing, and $(b)$ follows from $Z_{i,t} f_{X_t}\left(\beta(c_{(1)}) - c_i - Z_{i,t}\right)\mathbb{I}\left\{c_i > c_{(1)}\right\} \geq 0$ for any sample path. Since $(1 - \sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}), p)) = \mathbb{I}\left\{\beta(c_{(1)}) \geq p\right\}$, one has:

$$
\frac{\partial \pi_{i,t}(b_i, c_i, \boldsymbol{\beta})}{\partial c_i}|_{b_i=\beta(c_i)} \leq \mathbb{E}_{-i}\left[-r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t))\right]. \tag{C-6}
$$

Thus, applying the envelop theorem to $\pi_i$ yields:

$$
\pi_i(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}) - \pi_i(\beta(c_i), c_i, \boldsymbol{\beta}) \;=\; \int_{c_i}^{\bar{c}} \frac{\partial \pi_i(b_i, s, \boldsymbol{\beta})}{\partial s}|_{b_i=\beta(s)} ds \;=\; \int_{c_i}^{\bar{c}} \sum_{t=1}^{T} \frac{\partial \pi_{i,t}(b_i, s, \boldsymbol{\beta})}{\partial s}|_{b_i=\beta(s)} ds.
$$

Note that:

$$
\pi_i(\beta(c_i), c_i, \boldsymbol{\beta}) \;=\; \sum_{t=1}^{T}\mathbb{E}_{-i}\left[m_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t) - (c_i + X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t) + \left[p_{i,t} - c_i - X_t\right] A_i (1 - \sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t))\right].
$$

Combining the above two equations, one has:

$$
\begin{aligned}
\mathbb{E}_{-i}\left[\sum_{t=1}^{T} m_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t)\right] &= \sum_{t=1}^{T}\left[\pi_{i,t}(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}) - \int_{c_i}^{\bar{c}} \frac{\partial \pi_{i,t}(\beta(s), s, \boldsymbol{\beta})}{\partial s} ds\right] \\
&\quad + \sum_{t=1}^{T}\mathbb{E}_{-i}\left[(c_i + X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t) \;-\; \left[p_{i,t} - c_i - X_t\right] A_i (1 - \sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}), p_t))\right].
\end{aligned}
$$

If the auctioneer does not buy from one of the FA bidders in the FA, he buys from the spot market. Therefore, the expected buying price for Naive FA is:

$$
\begin{aligned}
\mathbb{E}[P^{NFA}] &= \mathbb{E}\left[\sum_{i=1}^{N}\sum_{t=1}^{T} m_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t) + \sum_{t=1}^{T}\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right) p_t(\boldsymbol{c},X_t,\boldsymbol{Z}_t)\right] \\
&= \sum_{t=1}^{T}\Bigg\{\mathbb{E}\left[\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right) p_t(\boldsymbol{c},X_t,\boldsymbol{Z}_t)\right] + \sum_{i=1}^{N}\pi_{i,t}(\beta(\bar{c}),\bar{c},\boldsymbol{\beta}) - \mathbb{E}\left[\sum_{i=1}^{N}\int_{c_i}^{\bar{c}} \frac{\partial \pi_{i,t}(\beta(s),s,\boldsymbol{\beta})}{\partial s} ds\right] \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{N}\left((c_i+X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t) - A_i\left(p_{i,t}-c_i-X_t\right)\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right)\right)\right]\Bigg\} \\
&\overset{(a)}{\geq} \sum_{t=1}^{T}\Bigg\{\mathbb{E}\left[\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right) p_t(\boldsymbol{c},X_t,\boldsymbol{Z}_t)\right] + \sum_{i=1}^{N}\pi_{i,t}(\beta(\bar{c}),\bar{c},\boldsymbol{\beta}) + \mathbb{E}\left[\sum_{i=1}^{N}\int_{c_i}^{\bar{c}} r_{i,t}(\beta(s),\beta(\boldsymbol{c}_{-i}),p_t) ds\right] \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{N}\left((c_i+X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t) - A_i\left(p_{i,t}-c_i-X_t\right)\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right)\right)\right]\Bigg\} \\
&\overset{(b)}{\geq} \sum_{t=1}^{T}\Bigg\{\mathbb{E}\left[\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right) p_t(\boldsymbol{c},X_t,\boldsymbol{Z}_t)\right] + \sum_{i=1}^{N}\pi_{i,t}(\beta(\bar{c}),\bar{c},\boldsymbol{\beta}) + \mathbb{E}\left[\sum_{i=1}^{N}\int_{c_i}^{\bar{c}} r_{i,t}(\beta(s),\beta(\boldsymbol{c}_{-i}),p_t) ds\right] \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{N}\left((c_i+X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t) - A_i\left(p_{i,t}-c_i-X_t\right)\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right)\right)\right]\Bigg\} \\
&\overset{(c)}{=} \sum_{t=1}^{T}\Bigg\{\mathbb{E}\left[\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right)\cdot\left(\sum_{j=N+1}^{M} A_j p_j(\boldsymbol{c},X_t,\boldsymbol{Z}_t)\right)\right] + \sum_{i=1}^{N}\pi_{i,t}(\beta(\bar{c}),\bar{c},\boldsymbol{\beta}) \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{N} A_i\left(c_i+X_t\right)\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right)\right] \\
&\quad + \mathbb{E}\left[\sum_{i=1}^{N}\left(r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\left(c_i+X_t\right) + \int_{c_i}^{\bar{c}} r_{i,t}(\beta(s),\beta(\boldsymbol{c}_{-i}),p_t) ds\right)\right]\Bigg\} \\
&\overset{(d)}{\geq} \sum_{t=1}^{T}\left\{\sum_{i=1}^{N}\pi_{i,t}(\beta(\bar{c}),\bar{c},\boldsymbol{\beta}) + \mathbb{E}\left[\widetilde{q}_0(\boldsymbol{c},X_t)\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\right)\right] + \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\left(v(c_i)+X_t\right)\right]\right\} \\
&= \sum_{t=1}^{T}\left\{\sum_{i=1}^{N}\pi_{i,t}(\beta(\bar{c}),\bar{c},\boldsymbol{\beta}) + \mathbb{E}\left[\widetilde{q}_0(\boldsymbol{c},X_t)\right] + \mathbb{E}\left[\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)\left(v(c_i)+X_t-\widetilde{q}_0(\boldsymbol{c},X_t)\right)\right]\right\},
\end{aligned}
$$

where (a) follows from (C-6), (b) follows from the fact that $A_i\left(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{b},p_i)\right) \leq A_i(1-r_{i,t}(\boldsymbol{b},p))$, (c) follows from the facts that $p_t(\boldsymbol{c},X_t,\boldsymbol{Z}_t)=\sum_{i=1}^{M} A_i p_i(\boldsymbol{c},X_t,\boldsymbol{Z}_t)$, and (d) follows from changing the order of integration, and the definition of $v(\cdot)$, and the fact that

$$
\begin{aligned}
&\mathbb{E}\left[(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)))\cdot\left(\sum_{j=N+1}^{M} A_j p_j(\boldsymbol{c},X_t,\boldsymbol{Z}_t)\right) + \sum_{i=1}^{N} A_i(c_i+X_t)(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t))\right] \\
=\ &\mathbb{E}\left[(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)))\cdot\left(\sum_{j=N+1}^{M} A_j(c_j+Z_{j,t}+X_t) + \sum_{i=1}^{N} A_i(c_i+X_t)\right)\right] \\
\geq\ &\mathbb{E}\left[(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)))\cdot\sum_{i=1}^{M} A_i(c_i+X_t)\right] = \mathbb{E}\left[(1-\sum_{j=1}^{N} r_{j,t}(\boldsymbol{\beta}(\boldsymbol{c}),p_t)))\cdot\widetilde{q}_0(\boldsymbol{c},X_t)\right],
\end{aligned}
$$

by the definition of $\widetilde{q}_0(\boldsymbol{c},X_t)=\sum_{i=1}^{M} A_i(c_i+X_t)$. This concludes the proof. □

**Proof of Proposition 12.** Bidder $i$'s total profit in a flexible FA is given by:

$$\begin{aligned}
\pi_i(\beta(c_i), c_i, \boldsymbol{\beta}_{-i}) &= \sum_{t=1}^{T} \pi_{i,t}(b_i, c_i, \boldsymbol{\beta}_{-i}), \qquad \text{where,} \\
\pi_{i,t}(b_i, c_i, \boldsymbol{\beta}_{-i}) &= \mathbb{E}\left[(b_i - c_i - X_t) \cdot \mathbb{I}\left\{b_i < \beta_j(c_j),\ j \neq i, b_i < Z_t + X_t\right\}\right] \\
&\quad + \mathbb{E}\left[(Z_t - c_i) \cdot \mathbb{I}\left\{b_i < \beta_j(c_j),\ j \neq i, b_i \geq Z_t + X_t \geq c_i + X_t\right\}\right].
\end{aligned}$$

Taking derivative with respect to $c_i$, one has

$$\begin{aligned}
\frac{\partial \pi_{i,t}(b_i, c_i, \boldsymbol{\beta})}{\partial c_i}\Big|_{b_i = \beta(c_i)} &= -\mathbb{E}\left[\mathbb{I}\left\{c_i < c_j,\ j \neq i, \beta(c_i) < Z_t + X_t\right\}\right] \\
&\quad + \mathbb{E}\left[\mathbb{I}\left\{c_i < c_j,\ j \neq i, \beta(c_i) \geq Z_t + X_t \geq c_i + X_t\right\}\right],
\end{aligned}$$

Define the following indicator random variables:

$$\begin{aligned}
n_i(c_i, \boldsymbol{b}, X_t, Z_t) &= \mathbb{I}\left\{b_i < b_j,\ j \neq i, b_i \geq Z_t + X_t \geq c_i + X_t\right\}, \\
r_i(\boldsymbol{b}, X_t, Z_t) &= \mathbb{I}\left\{b_i < b_j,\ j \neq i, b_i < Z_t + X_t\right\}.
\end{aligned}$$

Applying the Envelop Theorem to $\pi_i$ yields

$$\begin{aligned}
&\pi_i(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) - \pi_i(\beta(c_i), c_i, \boldsymbol{\beta}_{-i}) \\
&= \int_{c_i}^{\bar{c}} \sum_{t=1}^{T} \mathbb{E}_{-i}\left[-r_i(\beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) - n_i(s, \beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t)\right] ds,
\end{aligned}$$

Note that:

$$\pi_{i,t}(\beta(c_i), c_i, \boldsymbol{\beta}_{-i}) = \mathbb{E}_{-i}\Big[m_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) - (c_i + X_t) r_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) + [Z_t - c_i]\, n_i(c_i, \boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)\Big].$$

Combining the above two equations, and using the fact that $\pi_i(\beta(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) = 0$, one has

$$\begin{aligned}
\sum_{t=1}^{T} \mathbb{E}_{-i}\left[m_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)\right] &= \sum_{t=1}^{T} \Bigg\{ \mathbb{E}_{-i}\Big[(c_i + X_t) r_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) - [Z_t - c_i]\, n_i(c_i, \boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)\Big] \\
&\quad + \int_{c_i}^{\bar{c}} \mathbb{E}_{-i}\left[r_i(\beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) + n_i(s, \beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t)\right] ds \Bigg\}.
\end{aligned}$$

If the auctioneer does not buy from the FA winner, she buys from the spot market. Therefore, the expected buying price in the flexible FA is:

$$\mathbb{E}\left[P^{FLE}\right] = \mathbb{E}\left[\sum_{i=1}^{N} \sum_{t=1}^{T} m_i(\beta(\boldsymbol{c}), X_t, Z_t) + \sum_{t=1}^{T} \left(1 - \sum_{i=1}^{N} r_i(\beta(\boldsymbol{c}), X_t, Z_t)\right)(Z_t + X_t)\right] = \sum_{t=1}^{T} P_t^{FLE},$$

where

$$
\begin{aligned}
P_t^{FLE} \quad = \quad & \sum_{i=1}^{N} \mathbb{E}\left[(c_i + X_t) r_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) - [Z_t - c_i]\, n_i(c_i, \boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)\right] \\
& + \sum_{i=1}^{N} \mathbb{E}_{c_i}\left[\int_{c_i}^{\bar{c}} \mathbb{E}_{-i}\left[r_i(\beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) + n_i(s, \beta(s), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t)\right] ds\right] \\
& + \mathbb{E}\left[\left(1 - \sum_{i=1}^{N} r_i(\beta(\boldsymbol{c}), X_t, Z_t)\right)(Z_t + X_t)\right] \\
= \quad & \sum_{i=1}^{N} \mathbb{E}\left[(c_i + X_t) r_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) - [Z_t - c_i]\, n_i(c_i, \boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)\right] \\
& + \sum_{i=1}^{N} \mathbb{E}\left[\frac{F(c_i)}{f(c_i)} r_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) + \frac{F(c_i)}{f(c_i)} n_i(c_i, \boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)\right] + \mathbb{E}\left[\left(1 - \sum_{i=1}^{N} r_i(\beta(\boldsymbol{c}), X_t, Z_t)\right)(Z_t + X_t)\right] \\
\overset{(a)}{=} \quad & \mathbb{E}[X_t + Z_t] + \mathbb{E}\left[\sum_{i=1}^{N} (v(c_i) - Z_t) \cdot (r_i(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) + n_i(c_i, \boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t))\right] \\
\overset{(b)}{=} \quad & \mathbb{E}[X_t + Z_t] \\
& + \mathbb{E}\left[\sum_{i=1}^{N} (v(c_i) - Z_t) \cdot \mathbb{I}\{c_i < c_j,\ j \neq i\} \cdot (\mathbb{I}\{\beta(c_i) < Z_t + X_t\} + \mathbb{I}\{\beta(c_i) \geq Z_t + X_t \geq c_i + X_t\})\right],
\end{aligned}
$$

where (a) follows by the definition $v(c) = c + \frac{F(c)}{f(c)}$, and (b) follows from definition of $r_i$ and $n_i$. This concludes the proof. □

**Proof of Theorem 13.** In Proposition C4 in Appendix C.5.1, we provide expression for the expected price under FLR given by (see definitions of expressions also in Appendix C.5.1):

$$
\begin{aligned}
P^{FLR} \quad = \quad & \mathbb{E}_{x_0}\left\{\sum_{i=1}^{N}\sum_{t=1}^{T}(c_i + \frac{F(c_i)}{f(c_i)} - Z_t) r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - \sum_{i=1}^{N}\sum_{t=1}^{T} o_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) \frac{F(c_i)}{f(c_i)}\right\} \\
& + \sum_{t=1}^{T} \mathbb{E}_{x_0}[X_t + Z_t] \\
= \quad & \mathbb{E}_{x_0}\left\{\sum_{i=1}^{N}\sum_{t=1}^{T}(c_i + \frac{F(c_i)}{f(c_i)} - Z_t) \mathbb{I}\{c_i < c_j, j \neq i\} \cdot \tilde{r}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i)\right. \\
& \left. - \sum_{i=1}^{N}\sum_{t=1}^{T} \mathbb{I}\{c_i < c_j, j \neq i\} \cdot \tilde{o}_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) \frac{F(c_i)}{f(c_i)}\right\} + \sum_{t=1}^{T} \mathbb{E}_{x_0}[X_t + Z_t] \\
= \quad & \mathbb{E}_{x_0}\left\{\sum_{t=1}^{T}(v(c_{(1:N)}) - Z_t) \tilde{r}_{(1:N),t}\left(\beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)}\right)\right. \\
& \left. - \sum_{t=1}^{T} \tilde{o}_{(1:N),t}\left(\beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)}\right) \frac{F(c_{(1:N)})}{f(c_{(1:N)})}\right\} + \sum_{t=1}^{T} \mathbb{E}_{x_0}[X_t + Z_t]
\end{aligned}
$$

**First, we show that** $\lim_{N \longrightarrow \infty} \mathbb{E}_{x_0} \left[ \tilde{o}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \frac{F(c_{(1:N)})}{f(c_{(1:N)})} \right] = 0$ **for all** $t$**.** We show by induction (see proof at end of this proof) that there exists a finite $B_t$ such that

$$|\frac{\partial V_{t+1}(b_i, c_i, x_t)}{\partial c_i}| \leq B_t \qquad \text{for any } t. \tag{C-7}$$

As a result, one has

$$|\tilde{o}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i)| = \left| \frac{\partial V_{t+1}(X_t + Z_t, c_i, X_t) - V_{t+1}(b_i, c_i, X_t)}{\partial c_i} \cdot \mathbb{I}\{\text{Matching at } t\} \right| \leq 2B_t.$$

Since $c_{(1:N)} \longrightarrow \underline{c}$ in probability as $N$ goes to $\infty$, $\frac{F(c_{(1:N)})}{f(c_{(1:N)})} \longrightarrow 0$ in probability, by Bounded Convergence Theorem, $\lim_{N \longrightarrow \infty} \mathbb{E}_{x_0} \left[ \tilde{o}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \frac{F(c_{(1:N)})}{f(c_{(1:N)})} \right] = 0$ for any $t = 1, 2, \ldots, T$.

**Next, we show that**

$$\lim_{N \to \infty} \mathbb{E}_{x_0} \left[ (v(c_{(1:N)}) - Z_t) \tilde{r}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \right] = \mathbb{E}_{x_0} \left[ (\underline{c} - Z_t) \tilde{r}_{1,t} \left( \beta(\underline{c}), \mathbf{X}_t, \mathbf{Z}_t, \underline{c} \right) \right], \ \forall t.$$

Note that

$$\begin{aligned} & \mathbb{E}_{x_0} \left[ (v(c_{(1:N)}) - Z_t) \tilde{r}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \right] \\ = \ & \mathbb{E}_{x_0} \left[ (v(c_{(1:N)}) - \underline{c}) \tilde{r}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \right] \\ & + \mathbb{E}_{x_0} \left[ (\underline{c} - Z_t) \tilde{r}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \right] \end{aligned}$$

Since $v(c_{(1:N)}) \longrightarrow \underline{c}$ in probability as $N$ goes to $\infty$ and $|\tilde{r}_{(1:N),t}| \leq 2$, by Bounded Convergence Theorem, $\lim_{N \longrightarrow \infty} \mathbb{E}_{x_0} \left[ (v(c_{(1:N)}) - \underline{c}) \tilde{r}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \right] = 0$.

Note that $\tilde{r}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i) = \mathbb{I}\{A_t(b_i, c_i, X_t) \leq Z_t < b_i - X_t\} + \mathbb{I}\{b_i \leq X_t + Z_t\} = \mathbb{I}\{A_t(b_i, c_i, X_t) \leq Z_t \text{ or } b_i - X_t \leq Z_t\} \leq 1$, one gets:

$$\begin{aligned} & \mathbb{E}_{x_0} \left[ (\underline{c} - Z_t) \tilde{r}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \right] \\ & = \mathbb{E}_{x_0} \left[ (\underline{c} - Z_t) \mathbb{I}\{A_t(\beta(c_{(1:N)}), c_{(1:N)}, X_t) \leq Z_t \text{ or } b_i - X_t \leq Z_t\} \right]. \end{aligned}$$

Thus, one has

$$\lim_{N \to \infty} P^{FLR} = \sum_{t=1}^{T} \mathbb{E}[X_t + Z_t] + \sum_{t=1}^{T} \lim_{N \to \infty} \mathbb{E} \left[ (\underline{c} - Z_t) \, \tilde{r}_{(1:N),t} \left( \beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)} \right) \right].$$

Note that by the proof of Theorem 12, one has:

$$\lim_{N\to\infty} P^{FLE} = \sum_{t=1}^{T} \mathbb{E}[X_t + Z_t] + \sum_{t=1}^{T} \mathbb{E}\left[(\underline{c} - Z_t)\right].$$

The two previous expressions together yield:

$$\lim_{N\to\infty}[P^{FLE} - P^{FLR}] = \sum_{t=1}^{T} \lim_{N\to\infty} \mathbb{E}\left[(\underline{c} - Z_t)\cdot\left(1 - \tilde{r}_{(1:N),t}\left(\beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)}\right)\right)\right] \le 0.$$

Where the last inequality follows from the fact that $\underline{c} \le Z_t$ and $\tilde{r}_{(1:N),t}\left(\beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)}\right) \le 1$ almost surely (and noting that $\tilde{r}_{(1:N),t}\left(\beta(c_{(1:N)}), \mathbf{X}_t, \mathbf{Z}_t, c_{(1:N)}\right)$ is not equal to 1 almost surely, not even in the limit, because firms may not match in the restricted-flexible FA because of the dynamic incentives). □

**Proof of Equation** (C-7)**.** We show (C-7) by backward induction. Obviously, it holds for $t = T+1$ with $B_{T+1} = 0$ since $V_{T+1}(b, c, x) = 0$. Assume (C-7) holds for $t+1$, for some $t \in \{2, 3, \ldots, T+1\}$, next, we show it also holds for $t$. By (C-19), one has

$$\begin{aligned} \left|\frac{\partial V_t(b_i, c_i, x_{t-1})}{\partial c_i}\right| \;&\le\; \mathbb{E}_{x_{t-1}}\left[\left|\left(-1 + \frac{\partial V_{t+1}(X_t + Z_t, c_i, X_t) - \partial V_{t+1}(b_i, c_i, X_t)}{\partial c_i}\right)\right]\right| \\ &\quad + \mathbb{E}_{x_{t-1}}\left[\left|\frac{\partial V_{t+1}(b_i, c_i, X_t)}{\partial c_i}\right|\right] + 1 \\ &\le\; 2 + 3B_{t+1}. \end{aligned}$$

Thus, we have shown that (C-7) holds with $B_t = 2 + 3B_{t+1}$ for any $t$. □

**Proof of Proposition C1.** We show the ODE and boundary conditions for naive FA, monitored FA, flexible FA, and restricted-flexible FA separately in the follows.

**(1) ODE for Naive FA in diffused market.** First, we show that the bidding strategy satisfies the ODE equation. Taking derivative of profit function w.r.t $b$ and using the fact that $\frac{\partial \beta^{-1}(b)}{\partial b}|_{b=\beta(c)} = \frac{1}{\beta'(c)}$, one has

$$\begin{aligned} \frac{\partial \pi_i^{\text{NDFA}}(b, c, \beta)}{\partial b}\Big|_{b=\beta(c)} \;&=\; -\frac{(N-1)f(c)\bar{F}^{N-2}(c)}{\beta'(c)} \cdot \int_{\beta(c)-z_0}^{\bar{x}} \left(\beta(c) - c - x\right) f_X(x)dx \\ &\quad + \bar{F}^{N-1}(c)\cdot\left[\bar{F}_X\left(\beta(c) - z_0\right) - (z_0 - c) f_X(\beta(c) - z_0)\right]. \end{aligned}$$

Thus, by first-order-condition $\frac{\partial \pi_i^{\text{NDFA}}(b,c,\beta)}{\partial b} = 0$, we have shown that any symmetric equilibrium satisfies the ODE (C-6).

Next, we show the boundary condition. In the proof, we will omit the superscript "NDFA". Taking limit $c \nearrow z_0$ in the integral equation in Proposition 7, one obtains

$$\beta(z_0) \;=\; z_0 + \mathbb{E}\left[X \middle| \beta(z_0) \le z_0 + X\right].$$

It is easy to verify that the only solution to the above equation is $\beta(z_0) = z_0 + \bar{x}$ by applying L'Hôpital's rule. In what follows, we show $\beta'(z_0) = 0$. In (C-6), by taking limit $c \to z_0$, one obtains:

$$\begin{aligned}
\beta'(z_0) \;&=\; \lim_{c\to z_0}\left[\frac{(N-1)f(c)}{\bar{F}(c)} \cdot \frac{\mathbb{E}_X\left[(\beta(c)-c-X)\cdot \mathbb{I}\{\beta(c)\le z_0+X\}\right]}{\left[\bar{F}_X\left(\beta(c)-z_0\right)-(z_0-c)f_X(\beta(c)-z_0)\right]}\right] \\
&=\; \lim_{c\to z_0}\left[\frac{(N-1)f(c)\left[\int_{\beta(c)-z_0}^{\bar{x}}\left(\beta(c)-c-x\right)f_X(x)dx\right]}{\bar{F}(c)\left[\bar{F}_X\left(\beta(c)-z_0\right)-(z_0-c)f_X(\beta(c)-z_0)\right]}\right] \\
&=\; \frac{(N-1)f(z_0)}{\bar{F}(z_0)} \cdot \lim_{c\to z_0}\left[\frac{\int_{\beta(c)-z_0}^{\bar{x}}\left(\beta(c)-c-x\right)f_X(x)dx}{\bar{F}_X\left(\beta(c)-z_0\right)-(z_0-c)f_X(\beta(c)-z_0)}\right] \\
&\overset{(a)}{=}\; \frac{(N-1)f(z_0)}{\bar{F}(z_0)} \cdot \lim_{K\to\bar{x}} \frac{\int_K^{\bar{x}}(K-x)f_X(x)dx}{\bar{F}_X(K)} \\
&\overset{(b)}{=}\; \frac{(N-1)f(z_0)}{\bar{F}(z_0)} \lim_{K\to\bar{x}}\left[-\frac{\int_K^{\bar{x}} f_X(x)dx}{f_X(K)}\right] \;=\; 0,
\end{aligned}$$

where (a) holds since $\beta(z_0) = z_0 + \bar{x}$, and (b) is obtained by L'Hôpital's rule.

**(2) ODE for Monitored FA.** Recall that, bidder's profit is

$$\begin{aligned}
\pi_i(b,c,\beta) \;&=\; \mathbb{E}[\mathbb{I}\{b \le \beta(c_{(1:N),-i})\}\cdot(\min\{b, c+X+\Delta\}-c-X)] \\
&=\; \bar{F}^{N-1}(\beta^{-1}(b))\cdot\left[\Delta\mathbb{P}(b > c+X+\Delta) + \int_{b-c-\Delta}(b-c-x)f_X(x)dx\right].
\end{aligned}$$

Taking derivatives w.r.t $b$, one has

$$\begin{aligned}
\frac{\partial \pi_i(b,c,\beta)}{\partial b}|_{b=\beta(c)} \;&=\; -(N-1)\bar{F}^{N-2}(c)f(c)\cdot\left(\frac{\partial \beta^{-1}(b)}{\partial b}|_{b=\beta(c)}\right)\cdot \mathbb{E}_X\left[\min\{\beta(c)-c-X,\Delta\}\right] \\
&\quad +\bar{F}^{N-1}(c)\cdot\mathbb{P}(X \ge \beta(c)-c-\Delta).
\end{aligned}$$

Since $\frac{\partial \beta^{-1}(b)}{\partial b}|_{b=\beta(c)} = 1/\beta'(c)$, $\frac{\partial \pi_i(b,c,\beta)}{\partial b}|_{b=\beta(c)} = 0$ yields the ODE.

Next, we derive the boundary condition. From Proposition 9, the bidding strategy also satisfies integral equation, i.e., taking limit at $\bar{c}$, one has

$$\beta(\bar{c}) \;=\; \bar{c} + \mathbb{E}\left[X\middle|\beta(\bar{c}) \le \bar{c} + X + \Delta\right] - \Delta\frac{1-\mathbb{P}(\beta(\bar{c}) \le \bar{c} + X + \Delta)}{\mathbb{P}(\beta(\bar{c}) \le \bar{c} + X + \Delta)}.$$

Let $\beta(\bar{c}) = \bar{c} + \Delta + K$, obviously, $K \le \bar{x}$ since the above equation is not well-defined otherwise, then $K$ satisfies

$$K + \Delta = \mathbb{E}\left[X\middle|K \le X\right] - \Delta\frac{1-\mathbb{P}(X \ge K)}{\mathbb{P}(X \ge K)}.$$

Let

$$H(k) = \begin{cases} k + \Delta - \mathbb{E}[X], & k \le 0 \\ (k+\Delta)\bar{F}_X(k) - \int_k^{\bar{x}} x f_X(x)dx + \Delta \cdot F_X(k), & k \in [0, \bar{x}] \end{cases}$$

then $K \in [0, \bar{x}]$ is solution to $H(K) = 0$. Note that, $H'(k) = \bar{F}_X(k) - (k+\Delta)f_X(k) + kf_X(k) + \Delta f_X(k) = \bar{F}_X(k) > 0$, thus, $H(k)$ is increasing when $k \le \bar{x}$ and there is a unique solution to $H(K) = 0$. This completes the proof.

**(3) ODE for flexible FA.** Similar to the naive FA in diffused market, one could show that the bidding strategy under flexible FA satisfies the respective ODE and boundary conditions; we omit the details.

**(4) ODE for restricted-flexible FA.** The bidder's prifit expression is given in (C-13) in Appendix C.5, taking derivative w.r.t $b_i$, one gets

$$\begin{aligned}\frac{\partial \pi_i(b_i, c_i, x_0, \beta)}{\partial b_i} \;&=\; -(N-1)\overline{F}^{N-2}(\beta^{-1}(b_i))f(\beta^{-1}(b_i))\frac{\partial \beta^{-1}(b_i)}{\partial b_i} \cdot V_1(b_i, c_i, x_0) \\ &\quad + \overline{F}^{N-1}(\beta^{-1}(b_i)) \cdot \frac{\partial V_1(b_i, c_i, x_0)}{\partial b_i}.\end{aligned}$$

Since $\beta$ is an equilibrium, the FOC $\frac{\partial \pi_i(b_i,c_i,x_0,\beta)}{\partial b_i}\Big|_{b_i=\beta(c_i)} = 0$ implies that

$$0 = -(N-1)\overline{F}^{N-2}(c_i)f(c_i)\frac{1}{\beta'(c_i)} \cdot V_1(\beta(c_i), c_i, x_0) + \overline{F}^{N-1}(c_i) \cdot \frac{\partial V_1(b_i, c_i, x_0)}{\partial b_i}\Big|_{b_i=\beta(c_i)}.$$

Thus, the equilibrium satisfies the following ODE:

$$\beta'(c_i) = \frac{(N-1)f(c_i) \cdot V_1(\beta(c_i), c_i, x_0)}{\overline{F}(c_i) \cdot \frac{\partial V_1(b_i,c_i,x_0)}{\partial b_i}\Big|_{b_i=\beta(c_i)}}. \tag{C-8}$$

Note that for $T = 2$,

$$V_2(b_1, c) = \mathbb{E}\left[(b_1 - c - X_2)\mathbb{I}\{b_1 \le X_2 + z_0\} + (z_0 - c)\mathbb{I}\{b_1 > X_2 + z_0\}\right],$$

and the matching condition is equivalent to

$$V_2(b_0,c) \le (z_0 - c) + V_2(x_1 + z_0, c) \text{ and } b_0 > x_1 + z_0 \iff b_0 - z_0 > x_1 \ge B(b_0, z_0, c).$$

Then, we can simplify to:

$$\begin{aligned} V_1(b_0,c) &= \int_{B(b_0,z_0,c)}^{b_0-z_0} \left[(z_0 - c) + V_2(x+z_0,c) - V_2(b_0,c)\right] g(x)dx + V_2(b_0,c) \\ &\quad + \mathbb{E}\left[(b_0 - c - X_1)\,\mathbb{I}\{b_0 \le X_1 + z_0\}\right], \end{aligned}$$

to obtain the result. □

## C.4 Revenue Equivalence Between FPA without Reserve Price and FPA with Reserve Price

In this section we show that FPA without reserve prices is asymptotically equivalent to FPA with a random reserve price that equals to the spot market price determined as in the diffused naive FA.

Under the FPA with random reserve price (denoted as "RFPA"), the allocation function is

$$r_{i,t}(b_i, \boldsymbol{b}_{-i}, p) = \mathbb{I}\left\{b_i \le b_j, \forall j \ne i\right\} \mathbb{I}\left\{b_i \le p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t)\right\},$$

where $p_t(\boldsymbol{c}, X_t, \boldsymbol{Z}_t) = c + Z_t + X_t$ is spot market price under the diffused naive FA, and $(c, Z_t)$ is independent to $\{(c_i, Z_{i,t}), i = 1, 2, \ldots, N\}$, because of the diffused market assumption. However, note that $(c, Z_t)$ has the same marginal distribution to $(c_i, Z_{i,t})$. We have the following result.

**Proposition C2 (Revenue Equivalence)** *The expected buying price (measured at $t = 0$) for running the FPA without reserve price at every time period is given by:*

$$\mathbb{E}[P^{FPA}] = \sum_{t=1}^{T} \mathbb{E}\left[c_{(1:N)} + X_t + F(c_{(1:N)})/f(c_{(1:N)})\right]$$

*The expected buying price (measured at $t = 0$) for running the RFPA is given by:*

$$\mathbb{E}[P^{RFPA}] = \sum_{t=1}^{T} \mathbb{E}\left[\min\left\{c_{(1:N)} + F(c_{(1:N)})/f(c_{(1:N)}) + X_t,\ c + Z_t + X_t\right\}\right].$$

*In addition, we have*

$$\lim_{N\to\infty} \mathbb{E}[P^{FPA} - P^{RFPA}] = 0.$$

**Proof.** By standard arguments based on the envelope theorem, one can easily get the expressions for $\mathbb{E}[P^{FPA}]$ and for $\mathbb{E}[P^{RFPA}]$ [Milgrom, 2004]. Next, we show $\lim_{N\to\infty}\mathbb{E}[P^{FPA}-P^{RFPA}]=0$. Substituting the equations for $\mathbb{E}[P^{FPA}]$ and $\mathbb{E}[P^{RFPA}]$, one gets

$$\mathbb{E}[P^{FPA}-P^{RFPA}]=\sum_{t=1}^{T}\mathbb{E}\left[c_{(1:N)}+F(c_{(1:N)})/f(c_{(1:N)})-c-Z_t\right]^{+}.$$

Since $v(c_{(1:N)})=c_{(1:N)}+\frac{F(c_{(1:N)})}{f(c_{(1:N)})}$ converges to $\underline{c}$ in probability as $N\longrightarrow\infty$, one has

$$\lim_{N\to\infty}\mathbb{E}\left[c_{(1:N)}+F(c_{(1:N)})/f(c_{(1:N)})-c-Z_t\right]^{+}=0,$$

by the bounded convergence theorem. Thus, $\lim_{N\to\infty}\mathbb{E}[P^{FPA}-P^{RFPA}]=0$. □

## C.5 Analysis of Flexible-Restricted FA

We assume that $\{X_t : t\geq 0\}$ follows a Markov process. To simplify the initial analysis, we assume $T\geq 2$ periods and that $\{Z_t : t\geq 0\}$ is i.i.d.[1] We solve the model by backwards induction. The state variables are the current bid $b_t$, the previous realization of $x_{t-1}$, and the firm's cost $c$ (we ignore the subindex $i$ for now).

**t=T.** At the end of the horizon, the FA winner will match if and only if his cost is smaller than realized $z_T$. Hence, his expected payoff is given by:

$$V_T(b_T,c,x_{T-1})=\mathbb{E}_{x_{T-1}}\left[(b_T-c-X_T)\mathbb{I}\{b_T\leq X_T+Z_T\}+(Z_T-c)\mathbb{I}\{b_T>X_T+Z_T\geq X_T+c\}\right],\tag{C-9}$$

where the expectation over $X_T$ and $Z_T$ is taken conditional on the value $x_{T-1}$.

**t=1,..., T-1.** For a given realization of $X_t$ and $Z_t$, the bidder needs to decide whether to match or not. He faces a trade-off between making a sell today and decreasing the price for tomorrow. He solves the following optimization problem for realizations of $X_t$ and $Z_t$:

$$\widetilde{V}_t(b_t,c,x_t,z_t)\equiv\begin{cases}\max\{V_{t+1}(b_t,c,x_t),\ (z_t-c)+V_{t+1}(x_t+z_t,c,x_t)\}, & \text{if } b_t>x_t+z_t\\ b_t-c-x_t+V_{t+1}(b_t,c,x_t), & \text{o.w}\end{cases}\tag{C-10}$$

When $b_t>x_t+z_t$, a seller matches at period $t$ if and only if

$$V_{t+1}(b_t,c,x_t)\leq(z_t-c)+V_{t+1}(x_t+z_t,c,x_t).\tag{C-11}$$

[1] The assumption could be generalized into $\{(X_t,Z_t),t=0,1,\ldots,T\}$ is a Markov process with distribution of $(X_t,Z_t)$ depends on realization of $X_{t-1}$. To be consistent, we assume $X_t$ and $Z_t$ are independent conditional on $X_{t-1}$.

And $V_t(b_t, c, x_{t-1}) = \mathbb{E}_{x_{t-1}}\left[\widetilde{V}_t(b_t, c, X_t, Z_t)\right]$. In the following Proposition, we show that $V_t(b_t, c, x_{t-1})$ is strictly increasing in $b_t$.

**Proposition C3** *Assume that $X_t$ and $Z_t$ are independent for any given $x_{t-1}$, then, for any $t$:*

*(i) The value function $V_t(b_t, c, x_{t-1})$ is strictly increasing in $b_t$ for any $(c, x_{t-1})$ and $b_t < \bar{x} + \bar{z}$.*

*(ii) There exists $A_t(b_t, c, x_t)$ such that the FA winner matches spot market if and only if $A_t(b_t, c, x_t) \leq z_t < b_t - x_t$.*

**Proof.** We prove (i) by backward induction. First, we show that $V_T(b_T, c, x_{T-1})$ is strictly increasing in $b_T$. Note that

$$V_T(b_T, c, x_{T-1}) = \int_{\underline{z}}^{\bar{z}} \int_{b_T - z}^{\bar{x}} (b_T - c - x) g_{x_{T-1}}(x) dx f_{x_{T-1}}(z) dz + \int_c^{\bar{z}} (z - c) G_{x_{T-1}}(b_T - z) f_{x_{T-1}}(z) dz.$$

Where $g_{x_{T-1}}(\cdot), f_{x_{T-1}}(\cdot)$ are density function for $X_T, Z_T$, respectively, conditional on $X_{T-1} = x_{T-1}$. Taking derivative with respective to $b_T$, one gets

$$\begin{aligned}
&\frac{\partial V_T(b_T, c, x_{T-1})}{\partial b_T} \\
= \; & \mathbb{P}_{x_{T-1}}(b_T \leq X_T + Z_T) - \int_{\underline{z}}^{\bar{z}} (z - c) g_{x_{T-1}}(b_T - z) f_{x_{T-1}}(z) dz + \int_c^{\bar{z}} (z - c) g_{x_{T-1}}(b_T - z) f_{x_{T-1}}(z) dz \\
= \; & \mathbb{P}_{x_{T-1}}(b_T \leq X_T + Z_T) - \int_{\underline{z}}^{c} (z - c) g_{x_{T-1}}(b_T - z) f_{x_{T-1}}(z) dz \\
\geq \; & \mathbb{P}_{x_{T-1}}(b_T \leq X_T + Z_T) > 0.
\end{aligned}$$

Next, given that $V_{t+1}(b, c, x)$ is strictly increasing in $b$, we show that $V_t(b_t, c, x_{t-1})$ is strictly increasing in $b_t$. Since $V_{t+1}(b, c, x)$ is strictly increasing in $b$, by (C-11), there exists $A_t(b_t, c, x_t)$ such that matching if and only if $A_t(b_t, c, x_t) \leq z_t < b_t - x_t$.

In the following, we show that, $\widetilde{V}_t(b_t, c, x_t, z_t) \leq \widetilde{V}_t(b'_t, c, x_t, z_t)$ for any $b_t < b'_t$ and for any realization $(X_t, Z_t) = (x_t, z_t)$ and $c$, with inequality for samples that with a positive probability. We consider three cases: $(a) b_t < b'_t < z_t + x_t$, $(b) b_t < z_t + x_t \leq b'_t$, $(c) z_t + x_t \leq b_t < b'_t$.

(a) Case (a): $b_t < b'_t < z_t + x_t$. By (C-10), one has $\widetilde{V}_t(b_t, c, x_t, z_t) = b_t - c - x_t + V_{t+1}(b_t, c, x_t)$ and $\widetilde{V}_t(b'_t, c, x_t, z_t) = b'_t - c - x_t + V_{t+1}(b'_t, c, x_t)$. Since $V_{t+1}(b'_t, c, x_t) > V_{t+1}(b_t, c, x_t)$ and $b'_t > b_t$, obviously, $\widetilde{V}_t(b'_t, c, x_t, z_t) > \widetilde{V}_t(b_t, c, x_t, z_t)$.

(b) Case (b): $b_t < z_t + x_t \le b'_t$. By (C-10), one has $\widetilde{V}_t(b_t, c, x_t, z_t) = b_t - c - x_t + V_{t+1}(b_t, c, x_t)$ and $\widetilde{V}_t(b'_t, c, x_t, z_t) = \max\{V_{t+1}(b'_t, c, x_t),\ (z_t - c) + V_{t+1}(x_t + z_t, c, x_t)\} \ge (z_t - c) + V_{t+1}(x_t + z_t, c, x_t) > b_t - c - x_t + V_{t+1}(b_t, c, x_t)$, where the last inequality is because $z_t > b_t - x_t$. Therefore, one has $\widetilde{V}_t(b'_t, c, x_t, z_t) > \widetilde{V}_t(b_t, c, x_t, z_t)$.

(c) Case (c): $z_t + x_t \le b_t < b'_t$. By (C-10), one has $\widetilde{V}_t(b_t, c, x_t, z_t) = \max\{V_{t+1}(b_t, c, x_t),\ (z_t - c) + V_{t+1}(x_t + z_t, c, x_t)\}$ and $\widetilde{V}_t(b'_t, c, x_t, z_t) = \max\{V_{t+1}(b'_t, c, x_t),\ (z_t - c) + V_{t+1}(x_t + z_t, c, x_t)\} \ge \max\{V_{t+1}(b_t, c, x_t),\ (z_t - c) + V_{t+1}(x_t + z_t, c, x_t)\}$, where the last inequality is because $V_{t+1}(b'_t, c, x_t) \ge V_{t+1}(b_t, c, x_t)$. Therefore, one has $\widetilde{V}_t(b'_t, c, x_t, z_t) \ge \widetilde{V}_t(b_t, c, x_t, z_t)$.

Combining the above three cases, we have shown that $\widetilde{V}_t(b_t, c, x_t, z_t) \le \widetilde{V}_t(b'_t, c, x_t, z_t)$ for any sample paths and $\widetilde{V}_t(b_t, c, x_t, z_t) < \widetilde{V}_t(b'_t, c, x_t, z_t)$ when $z_t > b_t - x_t$. Since $\mathbb{P}_{x_{t-1}}(Z_t > b_t - X_t) > 0$ for any $b_t < \bar{z} + \bar{x}$, we have thus shown that $V_t(b'_t, c, x_{t-1}) = \mathbb{E}_{x_{t-1}}\left[\widetilde{V}_t(b'_t, c, X_t, X_t)\right] > \mathbb{E}_{x_{t-1}}\left[\widetilde{V}_t(b_t, c, X_t, X_t)\right] = V_t(b'_t, c, x_{t-1})$, namely, strictly monotonicity of $V_t(b_t, c, x_{t-1})$ in $b_t$.

The (ii) threshold result follows immediately from backward induction proofs in the part (i).

□

Note that: (a) The bidder matches spot market at period $T$ when $c \le z_T < b_T - x_T$, therefore, we denote $A_T(b_T, c, x_T) = c$ for notation convenience. (b) It is obvious from the definition that $A_t(b, c, x)$ is continuous in $b$ because of the continuity of $V_t(b, c, x)$ in $b$.

Thus, the FA winner's expected total profit is given by: for $t = 1, 2, \ldots, T-1$,

$$\begin{aligned}
& V_t(b_i, c, x_{t-1}) \\
= \ & \underbrace{\mathbb{E}_{x_{t-1}}\left[V_{t+1}(b_i, c, X_t)\mathbb{I}\{Z_t \le A_t(b_i, c, X_t)\}\right]}_{\text{(A) Not match spot market at } t} \\
& + \underbrace{\mathbb{E}_{x_{t-1}}\left[((Z_t - c) + V_{t+1}(X_t + Z_t, c, X_t)) \cdot \mathbb{I}\{A_t(b_i, c, X_t) \le Z_t < b_i - X_t\}\right]}_{\text{(B) Match spot market at } t} \\
& + \underbrace{\mathbb{E}_{x_{t-1}}\left[(b_i - c - X_t + V_{t+1}(b_i, c, X_t))\,\mathbb{I}\{b_i \le X_t + Z_t\}\right]}_{\text{(C) Beat spot market at } t} \\
= \ & \mathbb{E}_{x_{t-1}}\left[((Z_t - c) + V_{t+1}(X_t + Z_t, c, X_t) - V_{t+1}(b_i, c, X_t)) \cdot \underbrace{\mathbb{I}\{A_t(b_i, c, X_t) \le Z_t < b_i - X_t\}}_{\text{Matching spot market at } t}\right] \\
& + \mathbb{E}_{x_{t-1}}\left[V_{t+1}(b_i, c, X_t)\right] + \mathbb{E}_{x_{t-1}}\left[(b_i - c - X_t)\,\mathbb{I}\{b_i \le X_t + Z_t\}\right]. \qquad \text{(C-12)}
\end{aligned}$$

For each bidder $i$, given that his competitors play a strictly increasing strategy profile $\beta$, his profit with a bid $b_i$ is given by

$$\pi_i(b_i, c_i, x_0, \beta) = \mathsf{Pr}(b_i < \beta(c_j), \forall j \neq i) \cdot V_1(b_i, c_i, x_0) = \overline{F}^{N-1}(\beta^{-1}(b_i)) \cdot V_1(b_i, c_i, x_0). \quad \text{(C-13)}$$

Taking derivative with respect to $c_i$, one has

$$\frac{\partial \pi_i(b_i, c_i, x_0, \beta)}{\partial c_i}|_{b_i=\beta(c_i)} = \overline{F}^{N-1}(c_i) \cdot \frac{\partial V_1(b_i, c_i, x_0)}{\partial c_i}. \quad \text{(C-14)}$$

By Envelope Theorem,

$$\pi_i(\beta(\overline{c}), \overline{c}, x_0, \beta) - \pi_i(\beta(c_i), c_i, x_0, \beta) = \int_{c=c_i}^{\overline{c}} \overline{F}^{N-1}(c) \cdot \frac{\partial V_1(b_i, c, x_0)}{\partial c} dc.$$

Because $\pi_i(\beta(\overline{c}), \overline{c}, x_0, \beta) = 0$, one has

$$\pi_i(\beta(c_i), c_i, x_0, \beta) = -\int_{c=c_i}^{\overline{c}} \overline{F}^{N-1}(c) \cdot \frac{\partial V_1(b_i, c, x_0)}{\partial c}\Big|_{b_i=\beta(c)} dc. \quad \text{(C-15)}$$

### C.5.1 Mechanism Design Approach

Using direct revelation, a mechanism is characterized by $\{r_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i), m_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i), i = 1, 2, \ldots, N, t = 1, 2, \ldots, T.\}$, where $r_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i)$ specifies the allocation and $m_{i,t}$ specifies payment at period $t$, and $\mathbf{b} = (b_1, b_2, \ldots, b_N)$, $\mathbf{X}_t = (x_0, x_1, ..., x_t)$, and $\mathbf{Z}_t = (z_1, ..., z_t)$. In particular, for any $t = 1, 2, \ldots, T$, let

$$r_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i) = \mathbb{I}\{b_i < b_j, j \neq i\} \cdot \tilde{r}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i), \quad \text{(C-16)}$$

$$o_{i,t}(\mathbf{b}, X_t, Z_t, c_i) = \mathbb{I}\{b_i < b_j, j \neq i\} \cdot \tilde{o}_{i,t}(\mathbf{b}, X_t, Z_t, c_i). \quad \text{(C-17)}$$

Where

$$\tilde{r}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i) = \underbrace{\mathbb{I}\{A_t(b_i, c_i, X_t) \leq Z_t < b_i - X_t\}}_{\text{Match at } t} + \mathbb{I}\{b_i \leq X_t + Z_t\}, \quad \forall t = 1, 2, \ldots, T,$$

$$\tilde{o}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i) = \frac{\partial V_{t+1}(X_t + Z_t, c_i, X_t) - \partial V_{t+1}(b_i, c_i, X_t)}{\partial c_i} \cdot \mathbb{I}\{\text{Match at } t\}, \quad \forall t = 1, 2, \ldots, T-1,$$

and $\tilde{o}_{i,T}(\mathbf{b}, \mathbf{X}_T, \mathbf{Z}_T, c_i) = 0$. Intuitively speaking, $r_{i,t}$ gives allocation rule *ex-ante* for bidder $i$ and $\tilde{r}_{i,t}$ gives allocation rule *ex-post* allocation rule for bidder $i$ given he wins the FA at period 0.

Finally, the following proposition provides explicit expression of the buyer's expected total prices using mechanism design approach and Theorem 13 (see main text) compares the buyer's price under FLR and the one under FLE.

**Proposition C4** *The buyer's expected price under FLR is given by*

$$P^{FLR} = \mathbb{E}_{x_0}\Big\{\sum_{i=1}^{N}\sum_{t=1}^{T}(c_i + \frac{F(c_i)}{f(c_i)} - Z_t)r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - \sum_{i=1}^{N}\sum_{t=1}^{T}o_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)\frac{F(c_i)}{f(c_i)}\Big\} + \sum_{t=1}^{T}\mathbb{E}_{x_0}[X_t + Z_t]. \tag{C-18}$$

**Proof.** Bidder $i$'s total profit could also be represented by

$$\pi_i(\beta(c_i), c_i, x_0, \beta) = \mathbb{E}_{x_0}\left[\sum_{t=1}^{T}[m_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - (c_i + X_t)r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)]\right].$$

Therefore, substituting (C-15) into the above equation, one has

$$\begin{aligned}&\mathbb{E}_{x_0}\left[\sum_{t=1}^{T}m_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)\right]\\ &= \mathbb{E}_{x_0}\left[\sum_{t=1}^{T}(c_i + X_t)r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)\right] - \int_{c=c_i}^{\bar{c}}\overline{F}^{N-1}(c)\cdot\frac{\partial V_1(b_i, c, x_0)}{\partial c}\Big|_{b_i=\beta(c)}dc\end{aligned}$$

Buyer's total expected payment is given by

$$\begin{aligned}P^{FLR} &= \mathbb{E}_{x_0}\left[\sum_{t=1}^{T}\left[\sum_{i=1}^{N}m_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) + (Z_t + X_t)(1 - \sum_{i=1}^{N}r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i))\right]\right]\\ &= \mathbb{E}_{x_0}\Big\{\sum_{i=1}^{N}\sum_{t=1}^{T}(c_i + X_t)r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - \sum_{i=1}^{N}\int_{c=c_i}^{\bar{c}}\overline{F}^{N-1}(c)\cdot\frac{\partial V_1(b_i, c, x_0)}{\partial c}\Big|_{b_i=\beta(c)}dc\\ &\quad + \sum_{t=1}^{T}(Z_t + X_t)(1 - \sum_{i=1}^{N}r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i))\Big\}\\ &= \mathbb{E}_{x_0}\Big\{\sum_{i=1}^{N}\sum_{t=1}^{T}(c_i - Z_t)r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - \sum_{i=1}^{N}\int_{c=c_i}^{\bar{c}}\overline{F}^{N-1}(c)\cdot\frac{\partial V_1(b_i, c, x_0)}{\partial c}\Big|_{b_i=\beta(c)}dc\Big\}\\ &\quad + \sum_{t=1}^{T}\mathbb{E}_{x_{t-1}}[X_t + Z_t].\end{aligned}$$

Taking derivative w.r.t $c_i$ in (C-12), one gets

$$\begin{aligned}&\frac{\partial V_t(b_i, c_i, x_{t-1})}{\partial c_i}\\ &= \mathbb{E}_{x_{t-1}}\left[\left(-1 + \frac{\partial V_{t+1}(X_t + Z_t, c_i, X_t) - V_{t+1}(b_i, c_i, X_t)}{\partial c_i}\right)\cdot\mathbb{I}\{A_t(b_i, c, X_t) \le Z_t < b_i - X_t\}\right]\\ &\quad - \mathbb{E}_{x_{t-1}}\left[((A_t - c) + V_{t+1}(X_t + A_t, c, X_t) - V_{t+1}(b_i, c, X_t))\cdot f_{Z_t}(A_t)\cdot\frac{\partial A_t}{\partial c_i}\right]\\ &\quad + \mathbb{E}_{x_{t-1}}\left[\frac{\partial V_{t+1}(b_i, c_i, X_t)}{\partial c_i}\right] - \mathbb{E}_{x_{t-1}}\left[\mathbb{I}\{b_i \le X_t + Z_t\}\right].\end{aligned}$$

Since $A_t$ satisfies the matching equation (C-11), i.e., $V_{t+1}(b_i, c, X_t) = (A_t - c) + V_{t+1}(X_t + A_t, c, X_t)$. Thus,

$$\begin{aligned}
&\frac{\partial V_t(b_i, c_i, x_{t-1})}{\partial c_i} \\
= \; & \mathbb{E}_{x_{t-1}}\left[\left(-1 + \frac{\partial V_{t+1}(X_t + Z_t, c_i, X_t) - \partial V_{t+1}(b_i, c_i, X_t)}{\partial c_i}\right) \cdot \underbrace{\mathbb{I}\{A_t(b_i, c, X_t) \le Z_t < b_i - X_t\}}_{\text{Matching Event}}\right] \\
& + \mathbb{E}_{x_{t-1}}\left[\frac{\partial V_{t+1}(b_i, c_i, X_t)}{\partial c_i}\right] - \mathbb{E}_{x_{t-1}}\left[\mathbb{I}\{b_i \le X_t + Z_t\}\right]. \\
= \; & \mathbb{E}_{x_{t-1}}\left[-\tilde{r}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i) + \tilde{o}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i)\right] + \mathbb{E}_{x_{t-1}}\left[\frac{\partial V_{t+1}(b_i, c_i, X_t)}{\partial c_i}\right]. \qquad \text{(C-19)}
\end{aligned}$$

Recursively, one obtains

$$\frac{\partial V_1(b_i, c_i, x_0)}{\partial c_i} = \mathbb{E}_{x_0}\left[\sum_{t=1}^{T}\left[-\tilde{r}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i) + \tilde{o}_{i,t}(\mathbf{b}, \mathbf{X}_t, \mathbf{Z}_t, c_i)\right]\right].$$

Thus, one has

$$\begin{aligned}
&\overline{F}^{N-1}(c) \cdot \left.\frac{\partial V_1(b_i, c, x_0)}{\partial c}\right|_{b_i = \beta(c)} \\
= \; & \mathbb{E}_{x_0, \mathbf{c}_{-i}}\left[\mathbb{I}\{c < c_j, j \ne i\} \cdot \left.\frac{\partial V_1(b_i, c, x_0)}{\partial c}\right|_{b_i = \beta(c)}\right] \\
= \; & \mathbb{E}_{x_0, \mathbf{c}_{-i}}\left[\mathbb{I}\{c < c_j, j \ne i\} \cdot \left(\sum_{t=1}^{T}\left[-\tilde{r}_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) + \tilde{o}_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)\right]\right)\right] \\
= \; & \mathbb{E}_{x_0, \mathbf{c}_{-i}}\left[\sum_{t=1}^{T}\left[-r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) + o_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)\right]\right].
\end{aligned}$$

Thus, substituting the above equation into the expression of $P^{FLR}$, one has

$$\begin{aligned}
&P^{FLR} \\
= \quad & \mathbb{E}_{x_0}\Big\{ \sum_{i=1}^{N}\sum_{t=1}^{T}(c_i - Z_t) r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - \sum_{i=1}^{N}\int_{c=c_i}^{\bar{c}} (\sum_{t=1}^{T}[-r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) + o_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)])dc \Big\} \\
& + \sum_{t=1}^{T}\mathbb{E}_{x_0}[X_t + Z_t] \\
\overset{(a)}{=} \quad & \mathbb{E}_{x_0}\Big\{ \sum_{i=1}^{N}\sum_{t=1}^{T}(c_i - Z_t) r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - \sum_{i=1}^{N}(\sum_{t=1}^{T}[-r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) + o_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)]) \cdot \frac{F(c_i)}{f(c_i)} \Big\} \\
& + \sum_{t=1}^{T}\mathbb{E}_{x_0}[X_t + Z_t] \\
= \quad & \mathbb{E}_{x_{t-1}}\Big\{ \sum_{i=1}^{N}\sum_{t=1}^{T}(c_i + \frac{F(c_i)}{f(c_i)} - Z_t) r_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i) - \sum_{i=1}^{N}\sum_{t=1}^{T} o_{i,t}(\beta(\mathbf{c}), \mathbf{X}_t, \mathbf{Z}_t, c_i)\frac{F(c_i)}{f(c_i)} \Big\} \\
& + \sum_{t=1}^{T}\mathbb{E}_{x_0}[X_t + Z_t],
\end{aligned}$$

where (a) is because

$$\int_{\underline{c}}^{\bar{c}}\int_{c=c_i}^{\bar{c}} A(c)dcf(c_i)dc_i = \int_{\underline{c}}^{\bar{c}}\int_{\underline{c}}^{c} f(c_i)dc_i A(c)dc = \int_{\underline{c}}^{\bar{c}} F(c)A(c)dc = \mathbb{E}\left[A(c)\frac{F(c)}{f(c)}\right].$$

□

## C.6 The Optimal Mechanism

We consider the class of mechanisms defined in §3.5. The following analysis is similar to Milgrom [2004]. The following proposition characterizes the expected buying price in the BNE of a given FA.

**Proposition C5** *Let $\boldsymbol{\beta}(\cdot)$ be a BNE strategy profile induced by a mechanism $\boldsymbol{w} = (\boldsymbol{r}, \boldsymbol{m})$, such that equilibrium expected profits satisfy $\pi_i(\beta_i(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) = 0$, for all $i$. Then, the expected total buying price for the auctioneer is given by:*

$$\mathbb{E}[P] = \sum_{t=1}^{T}\mathbb{E}[Z_t + X_t] + \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)(v(c_i) - Z_t)\right], \tag{C-20}$$

*where the "virtual cost" function is $v(c) = c + F(c)/f(c)$.*

**Proof.** *Recall that if the auctioneer does not buy from one of the FA bidders, she buys from the spot market. Therefore, the total expected payments for an FA given its equilibrium strategy can be expressed as:*

$$\mathbb{E}[P] = \mathbb{E}\left\{\sum_{t=1}^{T}\sum_{i=1}^{N} m_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) + \sum_{t=1}^{T}\left(1 - \sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)\right)(Z_t + X_t)\right\}, \tag{C-21}$$

*where the expectation is taken with respect to the random variables $\boldsymbol{X} = (X_1, \ldots, X_T)$ and $\boldsymbol{Z} = (Z_1, \ldots, Z_T)$, and the random vector $\boldsymbol{c}$. In addition, throughout this proof we use the notation $\mathbb{E}_{-i,t}$ to denote expectation with respect to $X_t$, $Z_t$, and the random vector $\boldsymbol{c}_{-i}$. Consider the equilibrium payoff for bidder $i$:*

$$\pi_i(\beta_i(c_i), c_i, \boldsymbol{\beta}_{-i}) = \sum_{t=1}^{T}\mathbb{E}_{-i,t}\left[m_{i,t}(\beta_i(c_i), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) - (c_i + X_t) r_{i,t}(\beta_i(c_i), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t)\right]. \tag{C-22}$$

*Using the envelope theorem, and using the fact that* $\pi_i(\beta_i(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) = 0$*, we obtain:*

$$\pi_i(\beta_i(c_i), c_i, \boldsymbol{\beta}_{-i}) = \sum_{t=1}^{T} \int_{c_i}^{\bar{c}} \mathbb{E}_{-i,t} \left[ r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) \right] dy. \tag{C-23}$$

*Equating* (C-22) *and* (C-23) *we obtain:*

$$\begin{aligned} \sum_{t=1}^{T} \mathbb{E}_{-i} \left[ m_{i,t}(\beta_i(c_i), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) \right] &= \sum_{t=1}^{T} \mathbb{E}_{-i,t} \left[ (c_i + X_t) r_{i,t}(\beta_i(c_i), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) \right] \\ &+ \sum_{t=1}^{T} \int_{c_i}^{\bar{c}} \mathbb{E}_{-i,t} \left[ r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) \right] dy. \end{aligned}$$

*Replacing in equation* (C-21)*, using the fact that the private costs* $c_i$ *are independent across firms, we get:*

$$\begin{aligned} \mathbb{E}[P] &= \sum_{t=1}^{T} \mathbb{E} \left[ \sum_{i=1}^{N} \left[ (c_i + X_t) r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) + \int_{c_i}^{\bar{c}} r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) dy \right] \right] \\ &+ \sum_{t=1}^{T} \mathbb{E} \left[ \left( 1 - \sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t) \right) (Z_t + X_t) \right] \\ &= \sum_{t=1}^{T} \mathbb{E} \left[ \sum_{i=1}^{N} \left[ r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)(c_i - Z_t) + \int_{c_i}^{\bar{c}} r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X_t, Z_t) dy \right] \right] \\ &+ \sum_{t=1}^{T} E[Z_t + X_t] \end{aligned} \tag{C-24}$$

*Next, note that*

$$\begin{aligned} \mathbb{E} \left[ \int_{c_i}^{\bar{c}} r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X, Z) dy \right] &= \mathbb{E}_{-i} \left[ \int_{\underline{c}}^{\bar{c}} \int_{c_i}^{\bar{c}} r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X, Z) dy f(c_i) dc_i \right] \\ &= \mathbb{E}_{-i} \left[ \int_{\underline{c}}^{\bar{c}} \int_{\underline{c}}^{y} r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i}), X, Z) f(c_i) dc_i dy \right] \\ &= \mathbb{E}_{-i} \left[ \int_{\underline{c}}^{\bar{c}} r_{i,t}(\beta_i(y), \boldsymbol{\beta}_{-i}(\boldsymbol{c}_{-i})), X, Z)(F(y)/f(y)) f(y) dy \right] \\ &= \mathbb{E} \left[ r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X, Z) F(c_i)/f(c_i) \right], \end{aligned} \tag{C-25}$$

*where the first equation follows by the independence of the private costs and the second by changing the order of integration. Replacing* (C-25) *in* (C-24)*, we obtain:*

$$\mathbb{E}[P] = \sum_{t=1}^{T} \mathbb{E}[Z_t + X_t] + \mathbb{E} \left[ \sum_{t=1}^{T} \sum_{i=1}^{N} \left[ r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)(c_i + \frac{F(c_i)}{f(c_i)} - Z_t) \right] \right],$$

*proving the result.* □

We next consider the following structural assumption on the virtual cost.

**Assumption C1** *The virtual cost function* $v(c) = c + F(c)/f(c)$ *is strictly increasing in c, for all* $c \in [\underline{c}, \bar{c}]$.

In the following result we provide a characterization of mechanisms that minimize the expected buying price of the auctioneer.

**Proposition C6** *Suppose Assumption C1 holds. An augmented mechanism* $(\boldsymbol{w}, \boldsymbol{\beta})$ *minimizes the expected buying price for the auctioneer among all feasible augmented mechanisms if it satisfies* $\pi_i(\beta_i(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) = 0$, *for all* $i$, *and its allocation rule in period* $t = 1, 2, \ldots, T$ *under the BNE strategy profile* $\boldsymbol{\beta}$ *satisfies the following: (1) if* $v(c_{(1)}) \leq z_t$, *then buy from the lowest cost FA supplier; and (2) if* $v(c_{(1)}) > z_t$, *then buy from the spot market. Moreover, there exists at least one such augmented mechanism that achieves the optimum.*

**Proof.** *Following the same argument as Proposition C5, the expected buying price for an augmented feasible mechanism* $(\boldsymbol{w}, \boldsymbol{\beta})$ *satisfies:*

$$\begin{aligned}\mathbb{E}[P] &= \sum_{t=1}^{T} \mathbb{E}[Z_t + X_t] + \mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N} r_{i,t}(\boldsymbol{\beta}(\boldsymbol{c}), X_t, Z_t)(v(c_i) - Z_t)\right] + \sum_{i=1}^{N} \pi_i(\beta_i(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) \\ &\geq \sum_{t=1}^{T} \mathbb{E}[Z_t + X_t] + \mathbb{E}\left[\sum_{t=1}^{T} \min\left(0, \min_{i=1,\ldots,N}(v(c_i) - Z_t)\right)\right], \qquad \text{(C-26)}\end{aligned}$$

*where the inequality follows because for a feasible mechanism* $\pi_i(\beta_i(c_i), c_i, \boldsymbol{\beta}_{-i}) \geq 0$, *for all* $i$ *and* $c_i$, *and because* $\sum_{i=1}^{N} r_{i,t}(\boldsymbol{b}, x, z) \leq 1$ *and* $r_{i,t}(\boldsymbol{b}, x, z) \geq 0$, *for all* $\boldsymbol{b}, x, z$. *The right hand side of* (C-26) *provides a lower bound on the expected buying price for any feasible mechanism, therefore, a feasible mechanism that achieves it must be optimal. Hence, using the fact that* $v(\cdot)$ *is strictly increasing, a feasible augmented mechanism with an allocation rule in equilibrium like the one proposed in the statement of the proposition and that satisfies* $\pi_i(\beta_i(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) = 0$, *for all* $i$, *must be optimal.*

*To prove the second part of the proposition we construct a mechanism that achieves the optimum. Consider a "modified" second-price auction in which every period bidders submit bids* $b_i$ *and the spot market "submits" a bid equal to* $b_0^t = v^{-1}(z_t)$ *after observing the realization of* $Z_t = z_t$. *The lowest bid among* $b_0^t, b_1, \ldots, b_N$ *wins and sells the object. If one of the bidders* $1, \ldots, N$ *wins, after observing the realization of* $X_t$, *the auctioneer pays him* $b_{(2)} + x_t$, *where* $b_{(2)}$ *is the second lower order statistics among* $b_0^t, b_1, \ldots, b_N$. *Loosing bidders do not receive payments. Therefore, the actual payoff for a winning bidder* $i$ *is given by* $b_{(2)} + x_t - (c_i + x_t) = b_{(2)} - c_i$, *which for every period is the same as the payoff in a standard second price auction. Hence, truthful bidding is a dominant strategy, so that bidder* $i$ *submitting a bid* $b_i = c_i$ *is a BNE. Clearly, a bidder with cost* $\bar{c}$ *has no chance of winning and* $\pi_i(\beta_i(\bar{c}), \bar{c}, \boldsymbol{\beta}_{-i}) = 0$. *Moreover, the winning bidder is determined by the minimum between* $c_{(1)}$ *and* $v^{-1}(z_t)$. *Because* $v(\cdot)$ *is strictly increasing, it follows that the allocation rule satisfies: (1) if* $v(c_{(1)}) \leq z_t$, *then buy from the lowest cost FA supplier; and (2) if* $v(c_{(1)}) > z_t$, *then buy from the spot market. These facts prove the result.* □

## C.7 Existence of Equilibrium

We prove the existence of symmetric BNE for the naive FA with diffused markets and the monitored FA. We note that these results do not follow by standard existence results for first price auctions, because of the presence of the random common cost component and its correlation with the random spot market price.

**Proposition C7** *Assume that the action space* $\mathcal{A}$ *is restricted to be finite. (i) The naive FA game in diffused market admits a symmetric BNE in increasing strategies. (ii) The monitored FA game also admits a symmetric BNE in increasing strategies.*

We prove this result by applying and specializing the result of Athey [2001] that establishes the existence of BNE for a large class of games of incomplete information in two steps. For the corresponding game with a finite action space, the so-called *single-crossing condition (SCC)* is shown to be sufficient for the existence of an increasing and symmetric BNE. We note that while Proposition C7 assumes a single period, we can extend the result to multiple periods.[2] For games with continuous and compact action spaces (which include the settings discussed in the current paper) Athey [2001] also establishes the existence of a symmetric BNE by taking a limit of a sequence of games with finite action space as the granularity of the action space increases. In Appendix C.7.2 we extend Proposition C7 from a finite action set to a continuous and compact action space by showing that the regularity conditions required for the limiting argument are valid for naive FAs.

Now, we provide conditions under which BNE strategies are continuous and strictly increasing. We have the following result.

**Proposition C8** *Any symmetric and increasing BNE strategy must satisfy: (i) Under monitored FA, it is continuous and strictly increasing in $c \in [\underline{c}, \bar{c}]$; (ii) under naive FA with diffused market, it is strictly increasing in $c \in [\underline{c}, \bar{c}]$. Further,*
*if* $\operatorname{argmax}_{b\in\mathcal{A}} \mathbb{E}\left[(b-c-X)\mathbb{I}\left\{b \le X + c_j + Z_j\right\}\right]$ *is unique for all $c \in [\underline{c}, \bar{c}]$, it must also be continuous in $c \in [\underline{c}, \bar{c}]$.*

A sufficient condition for the argmax to be unique is that $\mathbb{E}\left[(b-c-X)\mathbb{I}\left\{b \le X + c_j + Z_j\right\}\right]$ is strictly quasi-concave in $b$, for all $c$. As an example, one can show this is the case when the common cost $X$ has a uniform distribution and $\mathbb{E}[X] - \underline{x} > \underline{z} + \underline{c}$.

Finally, we note that similar results to those alluded to in this section can be proved for the 'flexible FA'. For brevity, we will omit the proofs of the latter results.

### C.7.1 Proofs

**Proof of Proposition C7.** First, we introduce the following definition. A twice-differentiable function $h : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is called supermodular or log-supermodular, respectively, if for all $x$ and $\theta$:

$$\frac{\partial^2}{\partial x \partial \theta} h(x,\theta) \ge 0, \quad \text{or if } h > 0 \quad \frac{\partial^2}{\partial x \partial \theta} \ln\left(h(x,\theta)\right) \ge 0.$$

Note that these are sufficient conditions for supermodularity. There are weaker related conditions that do not require differentiability and use function differences for the case of discrete actions.

**Definition C1** *(Athey 2001) The Single Crossing Condition (SCC) for games of incomplete information is satisfied if for each $i = 1, 2, \ldots, N$, whenever every opponent $j \ne i$ uses a strategy $\beta_j$ that is increasing, player $i$'s profit function, $\pi_i(b_i, c_i, \boldsymbol{\beta}_{-i})$ is supermodular or log-supermodular in $(b_i, c_i)$.*

Now, we are ready to prove the proposition.

[2]For a multi-period model with i.i.d $\{X_t, Z_t\}$, the auction happens only once at the beginning of first period, and bidder considers total profit, instead of just one-period profit, when submitting his bid. Thus, the multi-period could be reduced into a one-period model.

**(i) Establishing SCC for monitored FA.** By Proposition 9, one has

$$\begin{aligned}\pi_i(b,c,\beta) &= \mathbb{P}[i\ wins\ with\ b]\cdot\mathbb{E}\left[\mathbb{I}\{b\le(c+X+Z_i)\}(b-c-X)\right] \\ &\quad +\mathbb{P}[i\ wins\ with\ b]\cdot\mathbb{E}\left[\mathbb{I}\{b>(c+X+Z_i)\}\left((c+X+Z_i)-c-X\right)\right].\end{aligned}$$

Taking derivatives with respect to $c$, one obtains:

$$\begin{aligned}\frac{\partial\pi_i(b,c,\beta)}{\partial c} &= \mathbb{P}[i\ wins\ with\ b]\cdot\mathbb{E}\left[-\mathbb{I}\{b\le(c+X+Z_i)\}\right] \\ &= -\mathbb{P}[i\ wins\ with\ b]\cdot\left[\mathbb{P}\left[b\le(c+X+Z_i)\right]\right]\end{aligned}$$

where $\mathbb{P}[i\ wins\ with\ b]$ is the probability bidder $i$ defeats its competitors' with a bid $b$:

$$\begin{aligned}&\mathbb{P}[i\ wins\ with\ b] \\ &= \underbrace{\mathbb{P}(\beta_j(c_j)>b,\ \text{for all } j\neq i)}_{\text{winning probability with no ties}} \\ &+ \underbrace{\sum_{k=1}^{N-1}\frac{\mathbb{P}(\text{exactly } k \text{ bidders other than } i \text{ bid } b \text{ and the rest higher than } b)}{k+1}}_{\text{winning probability with ties}}.\end{aligned}$$

It can be easily shown that the winning probability $\mathbb{P}[i\ wins\ with\ b]$ is decreasing in $b$, i.e., a higher bid induces a lower winning probability. Obviously, $\mathbb{P}\left[b\le(c+X+Z_i)\right]$ is decreasing in $b$ and $\mathbb{P}\left[b\le(c+X+Z_i)\right]\ge 0$. Thus, the partial derivative $\frac{\partial\pi_i(b,c,\beta_{-i})}{\partial c}$ is increasing in $b$. Therefore, by Definition C1, SCC is satisfied.

**(ii) Establishing SCC for naive FA in diffused market.** Recall that the private costs $\{c_j\}$ are i.i.d. and independent with the common cost $X$. For any increasing profile $\beta_{-i}$, we have

$$\pi_i(b,c,\beta_{-i}) = \mathbb{P}[i\ wins\ with\ b]\cdot\mathbb{E}\left[(b-c-X)\mathbb{I}\left\{b\le(X+c_j+Z_j)\right\}\right], \quad \text{(C-27)}$$

where $\mathbb{P}[i\ wins\ with\ b]$ is same as the one given above in (i). Thus, taking the partial derivatives with respect to $c$, we have

$$\frac{\partial\pi_i(b,c,\beta_{-i})}{\partial c} = -\mathbb{P}[i\ wins\ with\ b]\cdot\mathbb{P}(b\le(X+c_j+Z_j)).$$

Since both $\mathbb{P}(b\le(X+c_j+Z_j))$ and $\mathbb{P}[i\ wins\ with\ b]$ are decreasing in $b$, the partial derivative $\frac{\partial\pi_i(b,c,\beta_{-i})}{\partial c}$ is increasing in $b$. Therefore, by Definition C1, SCC is satisfied. The existence of a symmetric BNE in increasing strategies follows by Theorem 1 in Athey [2001]. □

## Proof of Proposition C8.

### (i) Strictly monotonicity and continuity for monitored FA.

- **<u>Part 1. Equilibrium is strictly increasing.</u>** By Section 3.3.2, one has

$$\pi_i(\beta(c),c,\beta) = \mathbb{P}[i\ wins\ with\ \beta(c)]\cdot\mathbb{E}\left[\min\{\beta(c),(c+X+Z)\}-c-X\right].$$

Now, we show that the equilibrium is strictly increasing by contradiction. Assume that there is an interval with positive length $[\hat{c}_1,\hat{c}_2]$, with $\hat{c}_2\le\bar{c}$, such that $\beta(c)=\hat{b}$ for all $c\in[\hat{c}_1,\hat{c}_2]$. We consider two cases. First, suppose that $\pi_i(\hat{b}, \hat{c}_2, \beta) > 0$. In this case, $\mathbb{E}\left[\min\{\hat{b}, (\hat{c}_2 + X + Z)\} - \hat{c}_2 - X\right] > 0$. It is simple to observe that the bidder with private cost $\hat{c}_2$ is strictly better off by unilaterally deviating from $\hat{b}$ to $\hat{b} - \delta$ for small enough $\delta > 0$. To see this, note that with this deviation, $\mathbb{P}[i\ wins\ with\ \beta(c)]$ increases by a strictly positive discrete amount and the second term $\mathbb{E}\left[\min\{\hat{b} - \delta, (\hat{c}_2 + X + Z)\} - \hat{c}_2 - X\right]$ remains essentially unchanged for small enough $\delta > 0$ by continuity.

The second case we consider is $\pi_i(\beta(\hat{c}_2), \hat{c}_2, \beta) = 0$. In this case, it must be that $\pi_i(\beta(c), c, \beta) > 0$, for $c \in [\hat{c}_1, \hat{c}_2)$, because the previous function is strictly decreasing in $c$ for $c \in [\hat{c}_1, \hat{c}_2)$. Then, we can repeat the previous argument.

- **Part 2. Equilibrium is continuous.** We show it by contradiction using the first part of the proof. Assume there is a symmetric and strictly increasing equilibrium $\beta(\cdot)$ and $\hat{c}_1 \in [\underline{c}, \bar{c}]$ such that $\beta(\cdot)$ has a jump at $\hat{c}_1$. Let the left-limit and right-limit of $\beta$ at $\hat{c}_1$ be $b_- = \lim_{c \nearrow \hat{c}_1} \beta(c)$ and $b_+ = \lim_{c \searrow \hat{c}_1} \beta(c)$, respectively. Then, $b_- < b_+$. By (C-27) and the fact that the ties happens with probability zero, one has: for any $b \in [b_-, b^+]$:

$$\pi_i(b, \hat{c}_1, \beta) \quad = \quad \bar{F}^{N-1}(\hat{c}_1) \cdot \mathbb{E}\left[\min\{b, (c + X + Z)\} - c - X\right] \tag{C-28}$$

This is impossible because $\mathbb{E}\left[\min\{b, (c + X + Z)\} - c - X\right]$ is strictly increasing in $b$. Thus, it must be continuous.

## (ii) Strictly monotonicity and continuity for naive FA.

- **Part 1. Equilibrium is strictly increasing.** By Section 3.3.1, one has

$$\pi_i(\beta(c), c, \beta) = \mathbb{P}[i\ wins\ with\ \beta(c)] \cdot \mathbb{E}\left[(b - c - X)\mathbb{I}\left\{\beta(c) \le c_j + Z_j + X\right\}\right].$$

Where $(c_j, Z_j)$ is independent with $c, X$. In the following, we let $Z = c_j + Z_j$. We need the following Lemma for the proof.

**Lemma C1** *Let $\beta(\cdot)$ be a symmetric and increasing BNE strategy of the naive FA model. Then, $\beta(c) < \bar{z} + \bar{x}$, for all $c < \bar{z}$.*

**Proof.** *We argue by contradiction. Suppose that $\beta(c) = \bar{z} + \bar{x}$, for some $c < \bar{z}$. Note that in this case, $\pi_i(\beta(c), c, \beta) = 0$, because bidder $i$ with private cost $c$ has no chance of defeating the spot market. We show that $\beta'(c) = \bar{z} + \bar{x} - \epsilon$, for small enough $\epsilon > 0$ is a profitable unilateral deviation, so the initially proposed strategy cannot be a BNE. Let $\{i\ wins\}$ be the event in which $\beta_i(c_i) \le \beta_j(c_j)$, $\forall j$ and bidder $i$ is selected in case of a tie. Then,*

$$\begin{aligned}
\pi_i(\beta'(c), c, \beta) &= \mathbb{P}[i\ wins] \cdot \mathbb{E}\left[(\beta'(c) - c - X) \cdot \mathbb{I}\left\{\beta'(c) \le Z + X\right\}\right] \\
&= \mathbb{P}[i\ wins] \cdot \mathbb{E}\left[((\bar{z} - c - \epsilon) + (\bar{x} - X)) \cdot \mathbb{I}\left\{\beta'(c) \le Z + X\right\}\right] \\
&= \mathbb{P}[i\ wins] \cdot \left[(\bar{z} - c - \epsilon)\, \mathbb{P}[\beta'(c) \le Z + X] + \mathbb{E}\left[(\bar{x} - X) \cdot \mathbb{I}\left\{\beta'(c) \le Z + X\right\}\right]\right].
\end{aligned}$$

*Clearly, $\mathbb{P}[i\ wins] > 0$, $\mathbb{P}[\beta'(c) \le Z + X] > 0$, and $(\bar{x} - x) \ge 0$, for all realizations $x$. Moreover, for small enough $\epsilon$, $\bar{z} - c - \epsilon > 0$. The result follows.* □

Now, we show that the equilibrium is strictly increasing by contradiction. Let us write:

$$\pi_i(\beta(c), c, \beta) = \mathbb{P}[i\ wins] \cdot \left(\beta(c) - c - \mathbb{E}\left[X \middle| \beta(c) \leq Z + X\right]\right) \mathbb{P}[\beta(c) \leq Z + X].$$

Assume that there is an interval with positive length $[\hat{c}_1, \hat{c}_2]$, with $\hat{c}_2 < \bar{z}$, such that $\beta(c) = \hat{b}$ for all $c \in [\hat{c}_1, \hat{c}_2]$. We consider two cases. First, suppose that $\pi_i(\beta(\hat{c}_2), \hat{c}_2, \beta) > 0$. In this case, $\left(\beta(\hat{c}_2) - \hat{c}_2 - \mathbb{E}_{Z,X}\left[X \middle| \beta(\hat{c}_2) \leq Z + X\right]\right) \mathbb{P}[\beta(\hat{c}_2) \leq Z + X] > 0$. It is simple to observe that the bidder with private cost $\hat{c}_2$ is strictly better off by unilaterally deviating from $\hat{b}$ to $\hat{b} - \delta$ for small enough $\delta > 0$. To see this, note that with this deviation, $\mathbb{P}[i\ wins]$ increases by a strictly positive discrete amount and the other terms in $\pi_i(\beta(\hat{c}_2), \hat{c}_2, \beta)$ remain essentially unchanged for small enough $\delta > 0$ by continuity.

The second case we consider is $\pi_i(\beta(\hat{c}_2), \hat{c}_2, \beta) = 0$. Because $\hat{c}_2 < \bar{z}$, $\beta(\cdot)$ is increasing, and Lemma C1, it must be that $\mathbb{P}[i\ wins] \cdot \mathbb{P}[\beta(\hat{c}_2) \leq Z + X] > 0$. Hence, it must be that $\beta(\hat{c}_2) - \hat{c}_2 - \mathbb{E}\left[X \middle| \beta(\hat{c}_2) \leq Z + X\right] = 0$. Take small enough $\epsilon$, for which $\beta(\hat{c}_2 - \epsilon) = \beta(\hat{c}_2)$. We have that $\pi_i(\beta(\hat{c}_2 - \epsilon), \hat{c}_2 - \epsilon, \beta) > 0$, and we can use the same argument regarding a unilateral deviation like in the first case. The result follows.

- **Part 2. Equilibrium is continuous.** We show it by contradiction using the first part of the proof. Assume there is a symmetric and strictly increasing equilibrium $\beta(\cdot)$ and $\hat{c}_1 \in [\underline{c}, \bar{z}]$ such that $\beta(\cdot)$ has a jump at $\hat{c}_1$. Let the left-limit and right-limit of $\beta$ at $\hat{c}_1$ be $b_- = \lim_{c \nearrow \hat{c}_1} \beta(c)$ and $b_+ = \lim_{c \searrow \hat{c}_1} \beta(c)$, respectively. Then, $b_- < b_+$. By (C-27) and the fact that the ties happens with probability zero, one has

$$\begin{aligned} \pi_i(b, \hat{c}_1, \beta) &= \mathbb{P}(b < \beta(c_j), j \neq i) \cdot \mathbb{E}\left[(b - \hat{c}_1 - X)\mathbb{I}\{b \leq X + Z\}\right] \\ &= \bar{F}^{N-1}(\hat{c}_1) \cdot \mathbb{E}\left[(b - \hat{c}_1 - X)\mathbb{I}\{b \leq X + Z\}\right] \quad \text{, for any } b \in [b_-, b^+]. \end{aligned} \tag{C-29}$$

  There are two cases to consider. Suppose $\beta(\hat{c}_1) = b_-$. Then, $b_-$ must be the maximum of $\mathbb{E}\left[(b - \hat{c}_1 - X)\mathbb{I}\{b \leq X + Z\}\right]$ by the previous equation. Moreover, by continuity and taking the limit $\lim_{c \searrow \hat{c}_1}$, $b_+$ must also be the maximum of $\mathbb{E}\left[(b - \hat{c}_1 - X)\mathbb{I}\{b \leq X + Z\}\right]$. This contradicts our assumption of the unique maximum of $\mathbb{E}\left[(b - c - X)\mathbb{I}\{b \leq X + Z\}\right]$ for any $c$. The second case is analogous, proving the result.

□

## C.7.2 Existence of Equilibrium for Compact Space

In this appendix, we extend the existence of the equilibrium under the naive FA with diffused market from a finite set (in Proposition C7) to a compact space. To that end, we need to verify that our basic FA game satisfies some technical conditions required for a limiting argument used by Athey [2001] to pass from games with finite action spaces to games with continuous action space. To simplify, we abuse notation and denote the sum of the random variables $c + Z$ as just $Z$.

Before presenting the existence proof, we present and prove the following lemma that we use below; we also referred to this result in the main body of the paper.

**Lemma C2** *For any random variable $Y$ with pdf $h(\cdot)$, cdf $H(\cdot)$ and support $[\underline{y}, \bar{y}]$ (possibly $\underline{y} = -\infty$ and/or $\bar{y} = \infty$), let*

$$B(a,b) = \mathbb{E}\left[Y \middle| a < Y < b\right] = \frac{\int_a^b t dH(t)}{H(b) - H(a)}, \quad \underline{y} \le a < b \le \bar{y}.$$

*Then, $B(a,b)$ is increasing in $a$ and $b$.*

**Proof.** For any $\underline{y} \le a < b \le \bar{y}$, we have

$$\frac{\partial B}{\partial a}(a,b) = \frac{-ah(a)[H(b) - H(a)] + h(a)\int_a^b t dH(t)}{[H(b) - H(a)]^2} = \frac{h(a)\int_a^b (t-a) dH(t)}{[H(b) - H(a)]^2} \ge 0,$$

$$\frac{\partial B}{\partial b}(a,b) = \frac{bh(b)[H(b) - H(a)] - h(b)\int_a^b t dH(t)}{[H(b) - H(a)]^2} = \frac{h(b)\int_a^b (b-t) dH(t)}{[H(b) - H(a)]^2} \ge 0.$$

Thus, we have proved the lemma. □

Next, we show that an increasing symmetric pure strategy BNE exists for the naive FA with diffused market, by applying Theorem 6 in Athey [2001]. For self-completeness, we briefly summarize notation and assumptions made in Theorem 6 of Athey [2001]. After introducing the theorem, we then show that all conditions are satisfied for the naive FA model.

**Part 1. Restatement of results in Athey [2001].** Consider a game of incomplete information between $I$ players, $i = 1, \dots, I$, where each player first observes his own type $t_i \in T_i = [\underline{t}_i, \bar{t}_i]$ and then takes an action $a_i$ from a compact set $\mathcal{A}_i \in \mathbb{R}$. Let $\mathcal{A} = \mathcal{A}_1 \times \cdots \times \mathcal{A}_I$, $\mathbf{T} = T_1 \times \cdots \times T_I$, $\underline{a}_i = \min \mathcal{A}_i$, and $\bar{a}_i = \max \mathcal{A}_i$. The joint density over player types is $f(\cdot)$, with the conditional density of $\mathbf{t}_{-i}$ given $t_i$ denoted $f(\mathbf{t}_{-i}|t_i)$. Player $i$'s payoff function is $u_i : \mathcal{A} \times \mathbf{T} \to \mathbb{R}$. Given any set of strategies for the opponents, $\alpha_j : T_j \to \mathcal{A}_j, j \ne i$, player $i$'s objective function is defined as follows (using the notation $(a_i, \alpha_{-i}(\mathbf{t}_{-i})) = (\cdots, \alpha_{i-1}(t_{i-1}), a_i, \alpha_{i+1}(t_{i+1}), \cdots)$):

$$U_i(a_i, t_i, \alpha_{-i}(\mathbf{t}_{-i})) = \int_{\mathbf{t}_{-i}} u_i\left((a_i, \alpha_{-i}(\mathbf{t}_{-i})), \mathbf{t}\right) f(\mathbf{t}_{-i}|t_i) d\mathbf{t}_{-i}.$$

**Assumption C2** *The types have joint density with respect to Lebesgue measure, $f(\cdot)$, which is bounded and atomless. Further, $\int_{\mathbf{t}_{-i} \in S} u_i\left((a_i, \alpha_{-i}(\mathbf{t}_{-i})), \mathbf{t}\right) f(\mathbf{t}_{-i}|t_i) d\mathbf{t}_{-i}$ exists and is finite for all convex $S$ and all increasing functions $\alpha_j : T_j \to \mathcal{A}_j, j \ne i$.*

For games with finite action spaces, say $\mathcal{A}_i = \{A_0, A_1, \dots, A_M\}$, she shows that the Kakutani's fixed point theorem is applicable when SCC is satisfied. Thus, a pure strategy BNE exists for games with finite action spaces. For games with compact action spaces, she assumes that player $i$'s payoff, given a realization of types and actions, has the following form

$$u_i(\mathbf{a}, \mathbf{t}) = \varphi_i(\mathbf{a}) \cdot \bar{v}_i(a_i, \mathbf{t}) + (1 - \varphi_i(\mathbf{a})) \cdot \underline{v}_i(a_i, \mathbf{t}) = \underline{v}_i(a_i, \mathbf{t}) + \varphi_i(\mathbf{a}) \cdot \Delta v_i(a_i, \mathbf{t}), \tag{C-30}$$

where $\Delta v_i(a_i, \mathbf{t}) = \bar{v}_i(a_i, \mathbf{t}) - \underline{v}_i(a_i, \mathbf{t})$. Intuitively, the winners receive payoffs $\bar{v}_i(a_i, \mathbf{t})$ with probability $\varphi_i(\mathbf{a})$, while losers receive payoffs $\underline{v}_i(a_i, \mathbf{t})$ with probability $1 - \varphi_i(\mathbf{a})$. In most auction models, participation is voluntary: there is some outside option such as not placing a bid that provides a fixed certain utility to the agent, typically normalized to zero. We refer to this action as $Q$. We introduce the following assumption.

**Assumption C3** *There exists $\lambda > 0$ such that, for all $i = 1, \ldots, I$, all $a_i \in [\underline{a}_i, \bar{a}_i]$ and all $\mathbf{t} \in \mathbf{T}$: (i) the types have support $T_1 \times \cdots \times T_I$; (ii) $\bar{v}_i(a_i, \mathbf{t})$ and $\underline{v}_i(a_i, \mathbf{t})$ are bounded and continuous in $(a_i, \mathbf{t})$; (iii) $\bar{v}_i(Q, \mathbf{t}) = 0, \underline{v}_i(Q, \mathbf{t}) = 0, \underline{v}_i(a_i, \mathbf{t}) \leq 0$, and $\Delta v_i(\bar{a}_i, \bar{\mathbf{t}}) < 0$; (iv) $\Delta v_i(a_i, \mathbf{t})$ is strictly increasing in $(-a_i, t_i)$; (v) for all $\varepsilon > 0$, $\Delta v_i(a_i, \mathbf{t}_{-i}, t_i + \varepsilon) - \Delta v_i(a_i, \mathbf{t}_{-i}, t_i) \geq \lambda\varepsilon$.*

Let $W_i(a_i, \alpha_{-i})$ denote the event that the realization of $\mathbf{t}_{-i}$ and the outcome of the tie-breaking mechanism are such that player $i$ wins with $a_i$, when opponents use strategies $\alpha_{-i}$ with the realization of $\mathbf{t}_{-i}$. Thus,

$$\mathbb{P}\left(W_i(a_i, \alpha_{-i}) | t_i\right) = \int \varphi_i\left(a_i, \alpha_{-i}(\mathbf{t}_{-i})\right) \cdot f(\mathbf{t}_{-i} | t_i) d\mathbf{t}_{-i}. \tag{C-31}$$

**Assumption C4** *For all $i = 1, \ldots, I$, all $a_i, a_i' \in [\underline{a}_i, \bar{a}_i]$, and whenever every opponent $j \neq i$ uses a strategy $\alpha_j$ that is increasing, $E_{\mathbf{t}_{-i}}\left[\Delta v_i(a_i, \mathbf{t}) \middle| t_i, W_i(a_i', \alpha_{-i})\right]$ is strictly increasing in $t_i$ and increasing in $a_i'$.*

**Theorem C1** (Athey 2001) *For all $i$, let $\mathcal{A}_i = Q \cup [\underline{a}_i, \bar{a}_i]$. Suppose Assumptios C2, C3, and C4 hold, and that the game satisfies the SCC. Then, there exists a pure strategy BNE in increasing strategies.*

It is simple to to use the previous result to establish the existence of a symmetric BNE for symmetric games with incomplete information, which is our case of interest.

**Part 2. Verifications of the Assumptions C2–C4.** Now, we are ready to show the existence of a BNE in the naive FA model by verifying the conditions in Assumptions C2-C4. This together with the verification of SCC guarantee the existence of increasing BNE. Our proof is presented for the general case of random $Z$. For this we need one additional assumption:

**Assumption C5** *Assume that the random variables $X$ and $Z$ satisfy: $\mathbb{E}[X | X + Z > b]$ is increasing in $b$, for all $b \geq 0$.*

The above assumption is used to guarantee Assumption C4. It can be shown that the condition in the assumption is satisfied in the following cases: 1) if $Z$ and $X$ are independent and identically distributed; and 2) if $Z$ and $X$ are both uniformly distributed (with potentially different supports). Since the bid has to be at least the private cost, thus the lowest possible rational bid is $\underline{c}$. For technical reasons, however, we define $\underline{b} = \underline{c} - \Delta,\ \Delta > 0$. Also, the bid will not be higher than $\bar{z} + \bar{x}$, namely, $\bar{b} = \bar{z} + \bar{x}$, which is the highest possible price in the spot market. To handle the two random components in the spot market price, we decompose the sport market into two "virtual" bidders: one has private cost $c_0^1 = x$ and the other has private cost $c_0^2 = z$, and they bid their true cost, i.e., $b_0^1 = x$ and $b_0^2 = z$. The FA winner competes with the "aggregate" price of these two virtual bidders, with bid $b_0 = b_0^1 + b_0^2$. To be consistent with notations in Athey [2001], we make the following transformation of the private cost and the bids: for all bidders $i = 1, 2, \ldots, N$ with private cost $c_i$ and bid $b_i$, let

$$\begin{aligned}
&a_i = \bar{c} + \bar{x} - b_i, \quad \underline{a}_i = \bar{c} + \bar{x} - \bar{b} = \bar{c} - z_0, \quad \bar{a}_i = \bar{c} + \bar{x} - \underline{b} = \bar{c} + \bar{x} - \underline{c} + \Delta, \\
&t_i = \bar{c} + \bar{x} - c_i, \quad \underline{t}_i = \bar{x}, \quad \bar{t}_i = \bar{c} + \bar{x} - \underline{c}, \\
&a_0^1 = t_0^1 = \bar{x} - x, \quad \underline{a}_0^1 = \underline{t}_0^1 = 0, \quad \bar{a}_0^1 = \bar{t}_0^1 = \bar{x} - \underline{x}, \\
&a_0^2 = t_0^2 = \bar{c} - z, \quad \underline{a}_0^2 = \underline{t}_0^2 = \bar{c} - \bar{z}, \quad \bar{a}_0^2 = \bar{t}_0^2 = \bar{c} - \underline{z}.
\end{aligned} \tag{C-32}$$

For any given $\mathbf{a} = (a_0^1, a_0^2, a_1, \ldots, a_N), \mathbf{t} = (t_0^1, t_0^2, t_1, \ldots, t_N)$, corresponding to (C-30), our naive FA model can be specified as follows: For any $i = 1, 2, \ldots, N$,

$$\varphi_i(\mathbf{a}) = \begin{cases} 1, & \text{if } b_i < b_j, \forall j \neq i \\ 0, & o.w. \end{cases} = \mathbb{I}\{b_i < b_j, \forall j \neq i\} = \mathbb{I}\{a_i > a_j, \forall j \neq i\}, \tag{C-33}$$

$$\underline{v}_i(a_i, \mathbf{t}) = 0, \quad \bar{v}_i(a_i, \mathbf{t}) = b_i - c_i - x = t_i - a_i - x = t_i - a_i - \bar{x} + t_0^1. \tag{C-34}$$

For simplification, we ignore ties in the winning probability in (C-33); a similar analysis applies if we consider them.

**Proposition C9** *Assume that Assumption C5 holds. Then, Assumptions C2-C4 hold for the naive FAs.*

**Proof.** We will check all conditions in Assumptions C2-C4 hold for naive FAs.

- Assumption C2: Assumption C2 is trivially true since $u_i(\mathbf{a}, \mathbf{t})$ is bounded for any $\mathbf{a}, \mathbf{t}$ by (C-30)-(C-34) and the fact that $a_i, t_i$ is bounded for any $i$.
- Assumption C3:
  - (i) and (ii) are trivial by (C-34) and the fact that $a_i, t_i$ is bounded for any $i$.
  - (iii). Let $Q = \bar{c} - \bar{z}$, i.e., the bid equals to the highest possible spot market price $\bar{z} + \bar{x}$. Thus, by bidding $Q$, the bidder will never win against spot market and $u_i(\mathbf{a}, \mathbf{t}|_{a_i = Q}) = 0$, thus, $a_i \geq Q$. $\Delta v_i(\bar{a}_i, \bar{\mathbf{t}}) = \bar{v}_i(\bar{a}_i, \bar{\mathbf{t}}) = \bar{t}_i - \bar{a}_i + \bar{t}_0^1 - \bar{x} = -\underline{x} - \Delta < 0$ by (C-32).
  - (iv). By (C-34), $\Delta v_i(a_i, \mathbf{t}) = \bar{v}_i(a_i, \mathbf{t}) = t_i - a_i + t_0^1 - \bar{x}$. Obviously, $\Delta v_i(a_i, \mathbf{t})$ is strictly increasing in $(-a_i, t_i)$.
  - (v). For all $\varepsilon > 0$, by (C-34), $\Delta v_i(a_i, \mathbf{t}_{-i}, t_i + \varepsilon) - \Delta v_i(a_i, \mathbf{t}_{-i}, t_i) = \varepsilon$. Thus, (v) in Assumption C3 is true for any $\lambda \in (0, 1]$.
- Assumption C4:

$$\begin{aligned}
\mathbb{E}_{\mathbf{t}_{-i}}\left[\Delta v_i(a_i, \mathbf{t}) \middle| t_i, W_i(a_i', \alpha_{-i})\right] &= \frac{\mathbb{E}_{\mathbf{t}_{-i}}\left[\Delta v_i(a_i, \mathbf{t}) \cdot \mathbb{I}\{W_i(a_i', \alpha_{-i})\} \middle| t_i\right]}{\mathbb{P}\left(W_i(a_i', \alpha_{-i}) \middle| t_i\right)} \\
&= \frac{\mathbb{E}_{\mathbf{t}_{-i}}\left[\left(t_i - a_i + t_0^1 - \bar{x}\right) \cdot \mathbb{I}\{W_i(a_i', \alpha_{-i})\} \middle| t_i\right]}{\mathbb{P}\left(W_i(a_i', \alpha_{-i}) \middle| t_i\right)} \\
&\overset{(a2)}{=} t_i - a_i - \bar{x} + \frac{\mathbb{E}_{\mathbf{t}_{-i}}\left[t_0^1 \cdot \mathbb{I}\{W_i(a_i', \alpha_{-i})\} \middle| t_i\right]}{\mathbb{P}\left(W_i(a_i', \alpha_{-i}) \middle| t_i\right)}
\end{aligned}$$

where (a2) holds because $\mathbf{t}_{-i}$ and $t_i$ are independent. By (C-31) and (C-33), we have

$$\begin{aligned}
&\frac{\mathbb{E}_{\mathbf{t}_{-i}}\left[t_0^1 \cdot \mathbb{I}\left\{W_i(a_i', \alpha_{-i})\right\} \middle| t_i\right]}{\mathbb{P}\left(W_i(a_i', \alpha_{-i}) \middle| t_i\right)} \\
= \quad & \frac{\int_{\mathbf{t}_{-i}} t_0^1 \cdot \mathbb{I}\left\{a_i' > \alpha_j(t_j), \forall j \neq i)\right\} f(\mathbf{t}_{-i}) d\mathbf{t}_{-i}}{\int_{\mathbf{t}_{-i}} \mathbb{I}\left\{a_i' > \alpha_j(t_j), \forall j \neq i)\right\} f(\mathbf{t}_{-i}) d\mathbf{t}_{-i}} \\
\stackrel{(a3)}{=} \quad & \frac{\mathbb{E}_{t_0^1, t_0^2}\left[t_0^1 \mathbb{I}\left\{a_i' > a_0^1 + a_0^2\right\}\right] \cdot \left(\prod_{j \neq i} \int_{\underline{t}_j}^{\alpha_j^{-1}(a_i')} f(t_j) dt_j\right)}{\mathbb{E}_{t_0^1, t_0^2}\left[\mathbb{I}\left\{a_i' > a_0^1 + a_0^2\right\}\right] \cdot \left(\prod_{j \neq i} \int_{\underline{t}_j}^{\alpha_j^{-1}(a_i')} f(t_j) dt_j\right)} \\
= \quad & \frac{\mathbb{E}_{t_0^1, t_0^2}\left[t_0^1 \mathbb{I}\left\{a_i' > a_0^1 + a_0^2\right\}\right]}{\mathbb{E}_{t_0^1, t_0^2}\left[\mathbb{I}\left\{a_i' > a_0^1 + a_0^2\right\}\right]} \\
\equiv \quad & \mathbb{E}\left[t_0^1 \middle| a_i' > t_0^1 + t_0^2\right],
\end{aligned}$$

where (a3) follows from the fact that $\mathbf{t}_{-i}$ are independent and the last equality from the facts that $a_0^1 = t_0^1$ and $a_0^2 = t_0^2$. Thus,

$$\mathbb{E}_{\mathbf{t}_{-i}}\left[\Delta v_i(a_i, \mathbf{t}) \middle| t_i, W_i(a_i', \alpha_{-i})\right] = t_i - a_i - \bar{x} + \mathbb{E}\left[t_0^1 \middle| a_i' > t_0^1 + t_0^2\right].$$

Obviously, $\mathbb{E}_{\mathbf{t}_{-i}}\left[\Delta v_i(a_i, \mathbf{t}) \middle| t_i, W_i(a_i', \alpha_{-i})\right]$ is strictly increasing in $t_i$. Next, we show that $\mathbb{E}\left[t_0^1 \middle| a_i' > t_0^1 + t_0^2\right]$ is increasing in $a_i'$. By tower property of conditional expectation and independence of $t_0^1$ and $t_0^2$, one has

$$\mathbb{E}_{t_0^1, t_0^2}\left[t_0^1 \middle| a_i' > t_0^1 + t_0^2\right] = \mathbb{E}\left[\bar{x} - X \middle| a_i' > \bar{x} - X + \bar{c} - Z\right] = \bar{x} - \mathbb{E}\left[X \middle| X + Z > \bar{x} + \bar{c} - a_i'\right].$$

The first equation follows by the definition $t_0^1 = \bar{x} - X$ and $t_0^2 = \bar{c} - Z$. Thus, by Assumption C5, one has $\mathbb{E}_{t_0^1, t_0^2}\left[t_0^1 \middle| a_i' > t_0^1 + t_0^2\right]$ is increasing in $a_i'$. Thus, Assumption C4 is true, namely, $\mathbb{E}_{\mathbf{t}_{-i}}\left[\Delta v_i(a_i, \mathbf{t}) \middle| t_i, W_i(a_i', \alpha_{-i})\right]$ is strictly increasing in $t_i$ and increasing in $a_i'$.

Therefore, we have shown conditions in Assumptions C2-C4 hold for the naive FAs. □

Recall that Athey's method is applicable to establishing the existence of a symmetric BNE when the game is symmetric. The above analysis establishes the conditions required to use Theorem C1 except SCC. This together with the SCC property established in Proposition C7 imply that a increasing symmetric BNE for naive FAs exists.

□

# Appendix D

# Appendix for Chapter 4

**Proof of Theorem 14.** While we provide the proof only for the $n$-supplier and $n$-retailer structure, the same approach applies to establish the existence and uniqueness of the equilibria for the other two structures. The function $d_i(\mathbf{w})$ is clearly linear in $\mathbf{w}$, decreasing in its own wholesale price $w_i$ and increasing in others' wholesale price $w_j$. By (4.7), for any $i = 1, 2, \ldots, n$ and $j \neq i$, we have

$$\frac{\partial d_i}{\partial w_i} = \frac{-1 - \sum_{k\neq i}^{n} \frac{\varphi_k}{1-\varphi_k}}{\lambda\left(\frac{1-\varphi_i}{\varphi_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\varphi_k}{1-\varphi_k}\right)} \quad \text{and} \quad \frac{\partial d_i}{\partial w_j} = \frac{\frac{\varphi_j}{1-\varphi_j}}{\lambda\left(\frac{1-\varphi_i}{\varphi_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\varphi_k}{1-\varphi_k}\right)}$$

as well as $\frac{\partial^2 d_i}{\partial w_i^2} = 0$ and $\frac{\partial^2 d_i}{\partial w_i \partial w_j} = 0$, where $\varphi_i = \frac{\lambda}{2\beta - v_i} \in (0, 1)$. Hence,

$$\begin{aligned} \frac{\partial d_i}{\partial w_i} + \sum_{j\neq i} \frac{\partial d_i}{\partial w_j} &= \frac{-1 - \sum_{k\neq i}^{n} \frac{\varphi_k}{1-\varphi_k} + \sum_{j\neq i} \frac{\varphi_j}{1-\varphi_j}}{\lambda\left(\frac{1-\varphi_i}{\varphi_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\varphi_k}{1-\varphi_k}\right)} \\ &= \frac{-1}{\lambda\left(\frac{1-\varphi_i}{\varphi_i}\right)\left(1 + \sum_{k=1}^{n} \frac{\varphi_k}{1-\varphi_k}\right)} < 0 \end{aligned}$$

or, equivalently,

$$\sum_{j\neq i} \frac{\partial d_i}{\partial w_j} < -\frac{\partial d_i}{\partial w_i}. \tag{D-1}$$

Taking the second-order derivatives in equation (4.8), we obtain

$$\frac{\partial^2 \pi_i^S}{\partial w_i^2} = \left(2 - v_i \frac{\partial d_i}{\partial w_i}\right)\frac{\partial d_i}{\partial w_i} \leq 0 \quad \text{and} \quad \frac{\partial^2 \pi_i^S}{\partial w_i \partial w_j} = \left(1 - v_i \frac{\partial d_i}{\partial w_i}\right)\frac{\partial d_i}{\partial w_j} \geq 0;$$

here, the inequalities follow from $2 - v_i \frac{\partial d_i}{\partial w_i} > 0$ and $1 - v_i \frac{\partial d_i}{\partial w_i} > 0$, as $\frac{\partial d_i}{\partial w_i} < 0$. Thus, we have a supermodular game for which an equilibrium exists.

Next we prove the equilibrium's uniqueness by showing that the Hessian matrix is strictly dominant:

$$\begin{aligned}\sum_{j\neq i}\left|\frac{\partial^2\pi_i^S}{\partial w_i\partial w_j}\right| &= \left(1-v_i\frac{\partial d_i}{\partial w_i}\right)\sum_{j\neq i}\frac{\partial d_i}{\partial w_j}\\ &\overset{(a)}{<} \left(1-v_i\frac{\partial d_i}{\partial w_i}\right)\cdot\left(-\frac{\partial d_i}{\partial w_i}\right)\\ &\overset{(b)}{\leq} \left(2-v_i\frac{\partial d_i}{\partial w_i}\right)\cdot\left(-\frac{\partial d_i}{\partial w_i}\right)\\ &= \left|\frac{\partial^2\pi_i^S}{\partial w_i^2}\right|.\end{aligned}$$

Here, the inequality (a) follows from $1-v_i\frac{\partial d_i}{\partial w_i}>0$ and (D-1) and the inequality (b) holds owing to the nonnegativity of $-\frac{\partial d_i}{\partial w_i}$. The game is therefore supermodular and its Hessian matrix is strictly dominant; hence, there exists a unique equilibrium to the suppliers' competition game (see Cachon and Lariviere 2005b). □

**Proof of Lemma 5.** While we provide the proof only for the $n$-supplier and $n$-retailer structure, the other two supply chain structures can be determined with the same approach. Both the existence and uniqueness follow directly from Theorem 14. When $\alpha_i - c_i = \xi$ for $i = 1, 2, \ldots, n$, the only equilibrium is symmetric. By solving the first-order condition $\frac{\partial \pi_i^S}{\partial w_i} = 0$ for $i = 1, 2, \ldots, n$, we obtain the solution for $w_i^*$ that is given in the first equation in (4.12). Substituting it into (4.7), we get $d^*$ in (4.12).

Next, we show that the equilibrium meets the individual rationality (IR) principle; in other words, the profits of the retailers and suppliers are all nonnegative. It is trivial that $d^* \geq 0$. We have

$$\begin{aligned}\pi_i^{S*}\geq 0 &\iff w_i^*-c_i-\frac{v}{2}d^*\geq 0\\ &\iff \frac{(2\beta-v)[1+(n-2)\varphi][1+(n-1)\varphi]+\frac{v}{2}[1+(n-2)\varphi]}{2\beta[2+(n-3)\varphi][1+(n-1)\varphi]-v\big[1+\varphi[2n-3+(n-3)(n-1)\varphi]\big]}\leq 1\\ &\iff 2\beta[1+(n-2)\varphi-(n-1)\varphi^2]-\frac{v}{2}[1+(n-2)\varphi-2(n-1)\varphi^2]\geq 0\\ &\iff \frac{4\beta}{v}\geq\frac{1+(n-2)\varphi-2(n-1)\varphi^2}{1+(n-2)\varphi-(n-1)\varphi^2},\end{aligned}$$

where the last inequality holds because $\frac{4\beta}{v} > 1 \geq \frac{1+(n-2)\varphi-2(n-1)\varphi^2}{1+(n-2)\varphi-(n-1)\varphi^2}$. Similarly, $\pi_i^{R*} \geq 0$ is equivalent to

$$\begin{aligned}&\alpha_i-w_i^*-\left(\beta-\frac{v}{2}\right)(1+2\varphi(n-1))d^*\geq 0\\ \iff\ &(2\beta-v)[1+(n-2)\varphi][1+(n-1)\varphi]\geq\left(\beta-\frac{v}{2}\right)(1+2\varphi(n-1))[1+(n-2)\varphi]\\ \iff\ &2+2\varphi(n-1)\geq 1+2\varphi(n-1),\end{aligned}$$

The last inequality is obviously true. Thus, we have proved the proposition. □

**Proof of Theorem 15.** (i) Proof of $n$-supplier and $n$-retailer structure. We assume without loss of generality that

$\beta = 1$. From Lemma 5, we obtain

$$\begin{aligned}\pi_i^{S*}(0) - \pi_i^{S*}(v) &= \frac{(2-\lambda)(2+\lambda(n-2))}{(4+\lambda(n-3))^2(2+\lambda(n-1))}\xi^2 \\ &\quad - \frac{(2-v+\lambda(n-2))\Big((4-v)(2-v)+\lambda(n-2)(4-v)-2\lambda^2(n-1)\Big)}{2\Big(8-6v+\lambda(-10+\lambda(n-3)(n-1)+6n)-\lambda(2n-3)v+v^2\Big)^2}\xi^2, \quad \text{(D-2)}\\ \pi_i^{R*}(0) - \pi_i^{R*}(v) &= \frac{(2+\lambda(n-2))^2}{(4+\lambda(n-3))^2(2+\lambda(n-1))^2}\xi^2 \\ &\quad - \frac{[2-v+\lambda(n-2)]^2(2-v)}{2\Big(8-6v+\lambda(-10+\lambda(n-3)(n-1)+6n)-\lambda(2n-3)v+v^2\Big)^2}\xi^2. \quad \text{(D-3)}\end{aligned}$$

In the following, we show the results of suppliers and retailers in the $n$-supplier and $n$-retailer structure in Parts 1 and 2, respectively.

**Part 1.** The proof for the supplier's preferences. By (D-2), $\pi_i^{S*}(0) - \pi_i^{S*}(v) \geq 0$ is equivalent to $G(\lambda) = G_0 + G_1\lambda + G_2\lambda^2 + G_3\lambda^3 + G_4\lambda^4 + G_5\lambda^5 \geq 0$, where

$$\begin{aligned}
G_0 &= -8(4-v)(2-v)^2,\\
G_1 &= 4(14-3v-n(6-v))(4-v)(2-v),\\
G_2 &= -600+2v(204-v39+2v^2)-(88-6(8-v)v)n^2-2n(-248+156v-26v^2+v^3),\\
G_3 &= -4(-99+n(119+(-37+n)n))-4(41+n(-44+n(10+n)))v+(15+n(-13+n+n^2))v^2,\\
G_4 &= 2\Big(-66+104n-43n^2-2n^3+3n^4-(n-2)^2(n^2+n-3)v\Big),\\
G_5 &= (n-3)(n-1)(n^3-n^2-4n+6).
\end{aligned}$$

It is easy to verify that $G(0) = G_0 < 0$ for any $n \geq 2$ and $v \leq 1$, and that $G(1) = 2 - 7n^2 - 5n^3 + n^4 + n^5 + (-4 + 8n + 14n^2 + 2n^3 - 2n^4)v + (1 - 9n - 5n^2 + n^3)v^2 + 2nv^3$. Because $G(1, v = 0) > 0$ and $G(1, v = 1) > 0$ for any $n \geq 2$, it could be verified that $G(1) > 0$ for any $n \geq 2$ and $v \in [0, 1]$, and that $G(\lambda)$ is quasi-convex over $[0, 1]$. Hence, there exists a $\lambda^S(n) \in (0, 1)$ such that $G(\lambda) \geq 0$ if and only if (iff) $\lambda \geq \lambda^S(n)$.

**Part 2.** The proof for the retailer's preferences. Similarly, $\pi_i^{R*}(0) - \pi_i^{R*}(v) \geq 0$ is equivalent to $H(\lambda) = H_0 + H_1\lambda + H_2\lambda^2 + H_3\lambda^3 + H_4\lambda^4 + H_5\lambda^5 + H_6\lambda^6 \geq 0$, where

$$\begin{aligned}
H_0 &= 8(2-v)^2v,\\
H_1 &= -8(2-v)[(n-2)v^2-2(n-3)v-4(n-1)],\\
H_2 &= 2\Big(8(n-1)(9n-23)-8(6-13n+4n^2)v-2(11-6n+n^2)v^2+(n-2)^2v^3\Big),\\
H_3 &= 4\Big(4n(83-47n+8n^2)-2n(69-41n+7n^2)v+(-3+(n-3)^2n)v^2-22(8-3v)\Big),\\
H_4 &= 4(n-1)\big(-151+217n-98n^2+14n^3\big)-2\big(87-224n+199n^2-72n^3+9n^4\big)v+(3-4n+n^2)^2v^2,\\
H_5 &= 2(n-3)(n-2)(n-1)(20+(n-4)n(6-v)-3v),\\
H_6 &= (-6+11n-6n^2+n^3)^2.
\end{aligned}$$

It is clear that $H(0) > 0$ for any $v > 0$. As in the preceding proof, it can be shown that $H(\lambda)$ is quasi-concave in $\lambda$ over $[0, 1]$. Moreover, $H(1) = (n-1)n(1+n)^2(n^2-n-4)-2(1+5n-n^2-6n^3-n^4+n^5)v+(1+4n-6n^2-4n^3+n^4)v^2+2n^2v^3$. As such if $H(1) \geq 0$ then $H(\lambda) \geq 0$ for any $\lambda \in [0, 1]$; otherwise, there exists a $\lambda^R(n) \in (0, 1)$ such that $H(\lambda) \geq 0$ iff $\lambda \leq \lambda^R(n)$.

By (D-3) to show that $\underline{\lambda}^R(n) = 1$ for any $n \geq 3$, it suffices to prove that $\bar{H} = H(\lambda = 1) = (n-1)n(1+n)^2(n^2 - n - 4) - 2(1 + 5n - n^2 - 6n^3 - n^4 + n^5)v + (1 + 4n - 6n^2 - 4n^3 + n^4)v^2 + 2n^2v^3 \geq 0$ for any $n \geq 3$. It is easy to verify that

$$\begin{aligned}
\bar{H}|_{v=0} &= (n-1)n(1+n)^2(n^2-n-4) \geq 0 \text{ for any } n \geq 3,\\
\bar{H}|_{v=1} &= (n^2-2n-1)(n^2-1)^2 \geq 0,\\
\frac{\partial \bar{H}}{\partial v}|_{v=0} &= -2(1 + n\big(5 + n[n(n-3)(2+n) - 1]\big)) \leq 0,\\
\frac{\partial \bar{H}}{\partial v}|_{v=1} &= -2n(n-1)(1+n)[(n-2)n-1] \leq 0
\end{aligned}$$

Because $\frac{\partial \bar{H}}{\partial v}(v)$ is a convex function, it follows that $\frac{\partial \bar{H}}{\partial v}(v) \leq 0$ for any $v \in [0, 1]$; hence, $\bar{H} \geq 0$ for any $v \in [0, 1]$. Therefore, by case (i), we have $\pi_i^{R*}(0) - \pi_i^{R*}(v) \geq 0$ for any $\lambda \in [0, 1]$ and $v \in [0, 1]$. Thus, the proposition is true with $\underline{\lambda}^R(n) = 1$.

(ii) Proof of $n$-supplier and 1-retailer structure. Similar to case (i), it could be shown that a quasi-convex function $\hat{G}(\lambda)$ and a quasi-concave function $\hat{H}(\lambda)$ exist such that (a) $\hat{G}(0) < 0 < \hat{H}(0)$; (b) $\hat{G}(\lambda)$ and $\hat{H}(\lambda)$ have at most one root on $[0, 1 - v/2)$ (recall that the discount and substitution rates have to satisfy $v < \bar{v} \equiv 2 - 2\lambda$, i.e., $\lambda < 1 - v/2$.), denoted as $\hat{\lambda}^S(n)$ and $\hat{\lambda}^R(n)$, respectively; and (b) $\hat{\pi}_i^{S*}(0) - \hat{\pi}_i^{S*}(v) \geq 0$ is equivalent to $\hat{G}(\lambda) \geq 0$ and $\hat{\pi}_i^{R*}(0) - \hat{\pi}_i^{R*}(v) \geq 0$ is equivalent to $\hat{H}(\lambda) \geq 0$. Therefore, the first part of the results could be shown in exactly the same way as case (i).

Next, we show that the threshold values satisfy $\hat{\lambda}^S(n) > \hat{\lambda}^R(n)$ for any $n \geq 2$. To this end, it suffices to show that $\hat{G}(\lambda) + \hat{H}(\lambda) < 0$ for any $\lambda \in [0, 1 - v/2)$ and that any $n \geq 2$. After some algebra, one obtains $\hat{G}(\lambda) + \hat{H}(\lambda) = -(2 + \lambda(n-3))v \cdot \hat{A}(\lambda)$, where $\hat{A}(\lambda) = \hat{A}_0 + \hat{A}_1\lambda + \hat{A}_2\lambda^2 + \hat{A}_3\lambda^3 + \hat{A}_4\lambda^4$ and

$$\begin{aligned}
&\hat{A}_0 = (2-v)^3, \quad \hat{A}_1 = (2-v)^2(n(7-v) - 13 + 2v),\\
&\hat{A}_2 = (2-v)[(18-5v)n^2 - (60-16v)n + 54 - 15v],\\
&\hat{A}_3 = 4\left[(5-2v)n^3 - (23-9v)n^2 + (38-15v)n - 22 + 9v\right],\\
&\hat{A}_4 = 4(n-3)(n-1)(n^2-2n+2).
\end{aligned}$$

Similar to case (i), it is not difficult to show that (a) $\hat{A}(0) = (2-v)^3 > 0$ and $\hat{A}(1 - \frac{v}{2}) = \frac{1}{4}(n-1)^2(n+1)[(2-v)n + v](2-v)^3 > 0$ for any $n \geq 2$ and $v \in [0, \bar{v})$; and (b) $\hat{A}(\lambda)$ is quasi-concave on the interval $[0, 1 - v/2)$. Thus, we show that $\hat{A}(\lambda) > 0$ for any $n \geq 2$, $v \in [0, \bar{v})$, and $\lambda \in [0, 1 - v/2)$. Combined with the fact that $(2 + \lambda(n-3)) > 0$ for any $n \geq 2$ and $\lambda \in [0, 1 - v/2)$, we thus show that $\hat{G}(\lambda) + \hat{H}(\lambda) < 0$ for any $n \geq 2$, $v \in [0, \bar{v})$, and $\lambda \in [0, 1 - v/2)$. This proves that $\hat{\lambda}^S(n) > \hat{\lambda}^R(n)$ for any $n \geq 2$ and $v \in [0, \bar{v})$.

(iii) The proof of the results for the 1-supplier and $n$-retailer structure could be shown along the same lines as the other two cases. □

**Proof of Theorem 16.** For any $\lambda > 0$ and $v > 0$, by Theorem 15, we have

$$\lim_{n\to\infty} n^2\pi_n^{S*}(0,\lambda) = \frac{2\beta-\lambda}{\lambda^2} > \frac{(2\beta-\lambda)-v/2}{\lambda^2} = \lim_{n\to\infty} n^2\pi_n^{S*}(v,\lambda),$$
$$\lim_{n\to\infty} n^2\pi_n^{R*}(0,\lambda) = \frac{\beta}{\lambda^2} > \frac{\beta-v/2}{\lambda^2} = \lim_{n\to\infty} n^2\pi_n^{R*}(v,\lambda),$$
$$\lim_{n\to\infty} n^2\hat{\pi}_n^{S*}(0,\lambda) = \frac{\beta-\lambda}{2\lambda^2} > \frac{(\beta-\lambda)-v/4}{2\lambda^2} = \lim_{n\to\infty} n^2\hat{\pi}_n^{S*}(v,\lambda).$$

This proves the theorem. □

**Proof of Theorem 17.** While we provide the proof only for the $n$-supplier and $n$-retailer structure, the same approach applies to the other two structures. From Table 4.1, it is easy to see that a supplier's equilibrium profit under a wholesale price–only contract is greater than under a quantity discount contract iff $\Delta\pi^S(\lambda) \equiv \lambda^2(n^2-5) + 12\beta\lambda - 8\beta^2 \geq 0$. Note that $\Delta\pi^S(0) = -8\beta^2 < 0$ and that $\Delta\pi^S(\beta) = (n^2-1)\beta^2 > 0$ for all $n \geq 2$. Two solutions for $\Delta\pi^S(\lambda) = 0$ are

$$\lambda_1 = \frac{-6+2\sqrt{2n^2-1}}{n^2-5}\beta \quad \text{and} \quad \lambda_2 = \frac{-6-2\sqrt{2n^2-1}}{n^2-5}\beta.$$

If $n = 2$, then $\lambda_1 = (6-2\sqrt{7})\beta = 0.7085\beta \in (0,\beta)$ and $\lambda_2 = (6+2\sqrt{7})\beta > \beta$. When $n \geq 3$, $\Delta\pi^S(\cdot)$ is convex and $\lambda_2 < 0$; therefore, $\Delta\pi^S(\cdot) \geq 0$ iff $\lambda \geq \frac{-6+2\sqrt{2n^2-1}}{n^2-5}\beta$.

Next, we show the results for the retailer. According to Table 4.1, it is easily verified that the retailer's equilibrium profit under a wholesale price–only contract is larger than under a quantity discount contract if and only if $\Delta\pi^R(\lambda) \equiv 4\beta^2 + \lambda^2(n-3)^2 + 2\beta\lambda(2n-7) \geq 0$. When $n \geq 4$, $\Delta\pi^R(\lambda) \geq 0$ because all of its three terms are nonnegative. If $n = 3$, then $\Delta\pi^R(\lambda) = 4\beta^2 - 2\beta\lambda = 2\beta(2\beta-\lambda) > 0$ for all $\lambda$. If $n = 2$, we have $\Delta\pi^R(\lambda) = 4\beta^2 + \lambda^2 - 6\beta\lambda \geq 0$ iff $\lambda \leq (3-\sqrt{5})\beta$. □