Credible Cheap-Talk Communication of Private Demand Information on Both the Forecast Average and Accuracy

Tao Lu† and Brian Tomlin‡

†School of Business, University of Connecticut, Storrs, CT 06269

‡Tuck School of Business, Dartmouth College, Hanover, NH 03755

July 18, 2024

Abstract

A canonical setting in supply chain research is one in which a retailer sources product, under a wholesale price contract, from a manufacturer that invests in capacity in advance of the retailer's order. When the retailer possesses private information about demand, it is well understood that credible cheap-talk communication is not possible absent considerations of trust or the reactive setting of the wholesale price. This understanding is based on the practice of demand information being shared as a point forecast. Motivated by the fact that some firms are now sharing information on forecast uncertainty along with the mean, we re-visit the canonical setting but allow the retailer to communicate its average demand and its forecast accuracy and allow the manufacturer to have multiple sources of capacity. We establish that credible and informative communication emerges in equilibrium under very general conditions. Moreover, when the manufacturer has multiple sources of capacity that differ in reservation and execution costs, the communication can be influential, strictly improve the manufacturer's expected profit, and result in a Pareto-improvement of supply chain profits. Our results suggest that both the forecast average and accuracy should be communicated in a supply chain not only because upstream firms benefit from a quantification of uncertainty but because communicating information about forecast accuracy (in addition to average demand) enhances the credibility of communication. We establish that improvements to the manufacturer's capacity portfolio (e.g., expansion or cost reduction) can hurt the manufacturer because of an associated reduction in information revelation. This negative effect can occur if the improvement alters or impacts the resource that, in isolation, provides the highest optimal service level.

Key words: supply chain management, multi-dimensional cheap talk, capacity procurement

1 Introduction

Investments in operations (capacity, inventory, etc.) are made under uncertainty and firms therefore engage in forecasting to reduce the uncertainty faced in their investment decisions. Downstream firms in a supply chain often possess better demand information than upstream firms. It has been long understood that, by sharing information, downstream firms can help upstream firms improve their forecasts and thus improve overall supply chain performance. However, because buying firms "routinely share demand forecasts but usually as single points" (Niranjan et al., 2022), it is well known (e.g., Cachon and Lariviere, 2001; Özer et al., 2011; Donohue et al., 2018) that downstream firm might self-interestedly exaggerate or "inflate forecasts to assure sufficient supply" (Terwiesch et al., 2005). This tendency can hurt upstream firms through overinvestment: for example, the contract manufacturer Solectron was left with \$4.7 billion of excess inventory after several large customers exaggerated their demand forecasts (Engardio, 2001). Not surprisingly, then, there is a credibility issue when a downstream supply-chain party seeks to communicate its demand forecast to an upstream party.[1]

In some circumstances, sophisticated supply chain contracts might be used by an upstream firm to elicit a downstream firm's truthful demand information or for a downstream firm to credibly signal its private forecast (e.g., Cachon and Lariviere, 2001; Özer and Wei, 2006; Taylor and Xiao, 2010; Hu et al., 2013; Feng et al., 2015). However, simple wholesale price contracts are widely used in many situations in which supply chain firms share demand information primarily through nonbinding and unverifiable forecasts (Cohen et al., 2003; Chu et al., 2017; Berman et al., 2019). Consider a single-period setting in which the downstream party ("the retailer") obtains a private point forecast (average demand) that it can communicate (through a nonbinding, costless and unverifiable message, i.e., cheap talk) to the upstream party ("the manufacturer") who then invests in a capacity quantity. The extant literature has established that in such a setting, credible communication is impossible in equilibrium because the retailer always has an incentive to exaggerate the estimated average demand to secure enough capacity, unless supply chain members have trust and trustworthy behaviors (Özer et al., 2011) or the manufacturer can set the wholesale price in response to the retailer's message (Chu et al., 2017).

Although the sharing of point forecasts is common in practice, there are examples in both capital-

[1]The sharing and exaggeration of point forecasts is also observed internally in firms when marketing is better informed about demand than manufacturing (e.g., Celikbas et al., 1999; Kolassa et al., 2023). Our focus, however, is on the communication between firms and not that internal to firms.

goods (Cohen et al., 2003) and consumer-goods (Niranjan et al., 2022) supply chains in which downstream firms share private multi-dimensional forecast information that quantifies, to some degree, the uncertainty in the forecast. The buyer in Cohen et al. (2003) began communicating an interval instead of a point forecast to enhance credibility. Walmart and Procter & Gamble share forecast scenarios with suppliers (Niranjan et al., 2022). Amazon shares forecast percentile information so that vendors can improve their inventory planning (Amazon Vendor Training Center, n.d.). In this paper, we re-examine the credible communication question in a manufacturer-retailer capacity investment setting where the retailer has multi-dimensional private information (average demand and forecast accuracy) and the manufacturer can invest in multiple sources of capacity that differ in their reservation and execution costs.

We establish that, different to the extant literature, informative and credible communication emerges in equilibrium under very general conditions. When the manufacturer has multiple sources of capacity that differ in reservation and execution costs, this informative communication can be influential, i.e., the manufacturer makes different investment choices under different retailer messages. Moreover, cheap-talk communication can strictly improve the manufacturer's expected profit and result in a Pareto-improvement of the supply chain parties' profits. Our results suggest that both the forecast average and accuracy should be communicated in a supply chain not only because upstream firms benefit from a quantification of forecast uncertainty but also because communicating information about the forecast accuracy (in addition to information about average demand) enhances the credibility of the communication. We observe that for symmetric demand distributions, the value of communication is highest when the manufacturer's optimal service level is 0.5, because this service-level environment induces more information revelation about the forecast accuracy, and accuracy information is particularly beneficial to the manufacturer. For left- and right-skewed distributions, the service levels leading to the highest value of communication are lower or higher than 0.5, respectively. We explore the impact of operational improvements on the manufacturer's portfolio of capacity sources (e.g., reducing the cost of an existing source or expanding the set of sources) and establish that such an improvement, even if free, may hurt the manufacturer (because of an associated reduction in information revelation) if the improvement alters or impacts the resource that, in isolation, provides the highest optimal service level. From a managerial perspective, therefore, cost reduction and capacity expansion opportunities need to be viewed not only through an operational lens but also through an informational lens.

We organize the rest of the paper as follows. We review the relevant literature in §2. We describe the model in §3 and present the results in §4. We discuss several model extensions in §5 and conclude the paper in §6. Proofs of all theoretical results and supplemental results referenced in the paper can be found in the Online Appendices.

2 Literature Review

Cheap-talk models have been used to study the communication of operationally-relevant information in a variety of important contexts, e.g., demand-forecast sharing in supply chains (Özer et al., 2011), delay announcements in service industries (Allon et al., 2011), inventory availability in retail settings (Allon and Bassamboo, 2011), new product launches (Berman et al., 2019), supply-chain social responsibility (Lu and Tomlin, 2022), project management (Beer and Qi, 2022), and manufacturing outsourcing (Lu, 2024).

The demand-forecast communication papers are those most relevant to our work. Özer et al. (2011) consider a supply chain in which a retailer sources a product from a manufacturer under an exogenous wholesale price contract.[2] The retailer privately obtains a forecast of average demand and then communicates this forecast to the manufacturer via cheap talk. The manufacturer in turn decides on a capacity level to build before demand is realized. Absent considerations of trust, Özer et al. (2011) establish that no informative or influential equilibrium exists because the retailer always prefers to over-report its demand forecast. However, based on behavioral experiments and analytical modeling, when notions of trust (on the part of the manufacturer) and trustworthiness (on the part of the retailer) are taken into account, Özer et al. (2011) establish that truthful information sharing can arise. Chu et al. (2017) explore a similar supply-chain setting but intentionally omit any behavioral concepts such as trust and trustworthiness. They establish that truthful communication can arise in equilibrium if the wholesale price is determined by the manufacturer after the communication stage. A key reason is that over-reporting demand gives the manufacturer an incentive to install a higher capacity level but also provides it with an incentive to charge the retailer a higher wholesale price. Due to this trade-off, the retailer might truthfully report its private demand forecast in equilibrium.

We consider a similar supply chain setting as that in Özer et al. (2011), but like Chu et al. (2017) we purposefully omit trust-based considerations. Motivated by the fact that many aspects of the retailer's forecasting process are private, we depart from these papers by exploring a setting in which

[2] Özer et al. (2011) refer to the downstream (upstream) party as the manufacturer (supplier), but when summarizing their work we refer to the downstream (upstream) party as the retailer (manufacturer) to be consistent with other cheap-talk papers, e.g., Chu et al. (2017), and with the terminology adopted in our paper.

the retailer has private information about both the average demand and the forecast accuracy. We also depart from these papers by allowing the manufacturer to have multiple sources of capacity with different levels of flexibility. We establish that even when the wholesale price is predetermined, as in Özer et al. (2011), communication on these two dimensions of private information can be truthful in equilibrium. Furthermore, the communication can be influential and strictly improve the manufacturer's payoff if the manufacturer has access to multiple capacity sources with different reservation and execution costs. Some other papers also examine cheap-talk communication of demand forecasts but focus on different factors that may lead to truthful communication such as repeated interactions (Ren et al., 2010), horizontal competition (Shamir and Shin, 2016) and partial vertical ownership (Avinadav and Shamir, 2023). Recently, Li et al. (2022) compare two behavioral economics theories, and with experimental data, they show that the trust-embedded model better explains truthful cheap talk than level-k bounded rationality.

Turning to the economics literature on cheap talk, pioneered by Crawford and Sobel (1982), our work is relevant to the studies on cheap talk games with multi-dimensional private information. Battaglini (2002) reveals the role of multiple senders in inducing informative communication. Chakraborty and Harbaugh (2007) show that credible information transmission is possible when a sender can communicate the ranking of the realized values of the multi-dimensional private information. Chakraborty and Harbaugh (2010) find that a sender with state-independent preferences can credibly communicate its private information with comparative statements. Our work, partly inspired by Chakraborty and Harbaugh (2010), is focused on a supply chain setting which differs from Chakraborty and Harbaugh (2010). In particular, in Chakraborty and Harbaugh (2010) the sender's utility is expressed as a function of the expected values of each private information dimension under the receiver's posterior belief. However, in our model the payoff of the sender (i.e., the retailer) depends on the total capacity level chosen by the receiver (i.e., manufacturer) which is the solution to a set of newsvendor-type optimality conditions, instead of a function of the expected values of private information. Moreover, the operational implications of communicating the average forecast and forecast accuracy, as highlighted by our paper, are not examined in the existing economics literature.

In our model, the manufacturer can procure (build) capacity from multiple sources that differ in their reservation and execution costs. The manufacturer's problem is thus related to the literature on capacity procurement with supply-option contracts. Martínez-de Albéniz and Simchi-Levi (2005) and Fu et al. (2010) study the optimization problem in which a firm facing uncertain demand chooses from a set of supply-option contracts. Each contract specifies a reservation and an execution cost. For each contract, the firm first decides on a quantity to reserve, and then based on realized demand, executes the contract up to the quantity reserved. Martinez-de Albeniz and Simchi-Levi (2009) examine the case when the firm's suppliers competitively determine their contract terms. Very different to this literature, our focus is on the cheap-talk communication of demand information between a downstream retailer and an upstream firm with multiple capacity options. Among other findings, our results highlight the implications of having multiple capacity sources on the value of communication.

3 Model

We consider a supply chain consisting of a manufacturer (she) and a retailer (he). The retailer sources a product from the manufacturer to sell at a per-unit price p in a single selling season with uncertain demand. Prior to the selling season, the retailer communicates his private demand forecast to the manufacturer via cheap talk, and the manufacturer in turn determines her capacity procurement (reservation). Then, after demand is realized, the manufacturer fills the retailer's demand up to the total capacity she procured earlier.

The Retailer's Demand Forecast. The market demand is given by $D = a + (1/b)\epsilon$, where a is the retailer's forecast of the average demand, ϵ is a random noise with mean zero, and b represents the accuracy of the retailer's forecast. A higher b signifies less noise in the retailer's forecast. For example, as b approaches infinity, the retailer's forecast a becomes perfectly accurate, i.e., the actual demand D will exactly match the forecast a at $b = \infty$. We assume that a is a continuous random variable over support $[\underline{a}, \bar{a}]$ with a cumulative distribution function (cdf) G and a probability density function (pdf) g, and b follows a two-point distribution with support $\{b_l, b_h\}$ where $0 < b_l < b_h$. Let ρ_t denote the probability of $b = b_t$ for each $t \in \{l, h\}$, where $\rho_h + \rho_l = 1$. The random noise ϵ has a cdf F. Ex ante, a and b are independent but may become dependent from the manufacturer's perspective after communication. In §5.2, we show that our main results can be generalized to the case when both a and b are continuous random variables.

The realizations of a and b are *both* the private information of the retailer. This reflects the reality that when forecasting demand, firms obtain a point estimate of the average demand (i.e., a) but they also have private knowledge about the accuracy of the estimated demand. Evaluating the forecast accuracy is one of the major steps in demand forecasting processes (see, e.g., Chapter

7 of Chopra, 2019). Because the forecasting process adopted and the amount of data available are privately known by the retailer, both the estimated average demand and the forecast accuracy level are the retailer's private information. We refer to the two-dimensional private information (a, b) as the retailer's forecast or type. Let $S = [\underline{a}, \bar{a}] \times \{b_h, b_l\}$ denote the space of the retailer's type. As will be shown, this two-dimensional private information plays a critical role in the possibility of informative communication.

The Manufacturer's Capacity Decision. The manufacturer can procure (reserve, build) capacity from two sources, indexed by $i = 1, 2$. For each capacity source i, let r_i denote the reservation cost per unit, and c_i denote the execution cost per unit. Without loss of generality, we assume $0 < r_1 < r_2$ and $c_1 > c_2 \geq 0$, and so the first capacity source is more flexible in the sense that it has a lower cost to reserve (but a higher cost to execute). We will refer to capacity source 1 (resp. choice 2) as the more flexible (resp. less flexible) capacity. In practice, it is common that a manufacturer has multiple sources of capacity with different levels of flexibility. For instance, in the electronics and fashion industries, a manufacturer often procures capacity from multiple suppliers (e.g., contract manufacturers) that offer different supply-option contracts in which different reservation and execution costs are specified; see, for example, Martínez-de Albéniz and Simchi-Levi (2005), Martinez-de Albeniz and Simchi-Levi (2009) and Fu et al. (2010). In such cases, the (r_i, c_i)'s in our model represent the reservation and execution costs offered in each supply-option contract. If the manufacturer produces in-house, then the (r_i, c_i)'s represent the cost structures of different production lines or facilities. For example, a highly-automated line would incur significant build (reservation) cost as compared to a manual line with flexibly-schedulable workers (that can scale with demand) but the automated line's marginal production (execution) cost would be lower. As is common in the cheap-talk demand-forecasting literature (e.g., Özer et al., 2011; Chu et al., 2017), we assume that the manufacturer's cost parameters are common knowledge so as to focus on the strategic communication of asymmetric demand information.

Following a standard assumption in the operations cheap-talk literature (e.g., Özer et al., 2011; Berman et al., 2019; Lu and Tomlin, 2022), we assume that the manufacturer and retailer trade based on a simple price-only contract with a given wholesale price w. We make this assumption not only because price-only contracts are prevalent in practice but also because such an assumption enables us to abstract away from other factors (e.g., pricing and mechanism design) that may drive informative communication. In particular, Chu et al. (2017) show that when only the average demand information is private, truthful and influential communication can arise in equilibrium if the wholesale price is endogenously determined by the manufacturer in response to the retailer's message, but that no informative equilibrium exists if the wholesale price is fixed (or equivalently, set before the communication stage). In this paper, we intentionally assume the wholesale price to be pre-determined so as to rule out the effect of the manufacturer's responsive pricing. To exclude uninteresting cases, assume $p > w$ and $w > r_i + c_i$ for all i such that it is profitable for the retailer to source product and for the manufacturer to use either capacity source. As is common, the manufacturer earns the same per-unit revenue (i.e., the wholesale price w in our model) from selling its product to the retailer regardless of which capacity source it uses (Martínez-de Albéniz and Simchi-Levi, 2005; Dada et al., 2007; Ang et al., 2017; Dong et al., 2022).

The manufacturer's problem consists of two stages: capacity reservation before demand realization, and execution after demand realization. Let K_i denote the amount of capacity that the manufacturer chooses to reserve from source i. In the capacity reservation stage, the manufacturer determines (K_1, K_2) to maximize the following expected profit

$$\Pi_m(K_1, K_2|m) = \mathbb{E}[R(K_1, K_2, D)|m] - \sum_{i=1}^{2} r_i K_i, \tag{1}$$

where the expectation is taken over $D = a + (1/b)\epsilon$ based on the manufacturer's posterior belief about (a, b) given any message m, and $R(K_1, K_2, D)$ is the optimal objective value of the second-stage execution problem as defined below.

$$\begin{aligned} R(K_1, K_2, D) = \quad \max_{x_1,x_2} \quad & w \min(x_1 + x_2, D) - \textstyle\sum_{i=1}^{2} c_i x_i \\ s.t. \quad & 0 \le x_i \le K_i \text{ for } i = 1, 2. \end{aligned} \tag{2}$$

In the execution problem (2), the manufacturer chooses production quantities (x_1, x_2) to maximize her profit given any demand realization D and capacity levels (K_1, K_2). The constraints in (2) ensure that each production quantity cannot exceed the corresponding capacity level reserved in the first stage.

The Retailer's Payoff. It is easy to show that the total production quantity delivered by the manufacturer will be $x_1^* + x_2^* = \min(K_1 + K_2, D)$, i.e., the manufacturer will either fulfill all the demand if the total capacity level is sufficient, or exhaust all the capacity otherwise. Therefore, the retailer's expected profit is a function of $K_1 + K_2$ as given below:

$$\Pi_r(K_1, K_2) = (p - w)\mathbb{E}[\min(D, K_1 + K_2)]. \tag{3}$$

Clearly, the retailer's payoff depends on the manufacturer's capacity levels only through the total capacity level $K_1 + K_2$, and moreover, it is increasing in $K_1 + K_2$. Throughout the paper, we use "increasing" and "decreasing" in the weak sense.

Equilibrium Concept. As is common in the literature, we adopt the Perfect Bayesian Equilibrium (PBE) as our solution concept. Before formally defining the PBE for our problem, we provide a sketch of the sequence of events: (i) The retailer first observes a private forecast, i.e., the realization of (a, b); (ii) the retailer communicates with the manufacturer by sending a costless and unverifiable message $m \in M$ where M is the message space, and the manufacturer forms a posterior belief, denoted by $\mu(a, b|m)$, about the retailer's forecast; (iii) based on her posterior belief, the manufacturer determines the optimal capacity levels (K_1^*, K_2^*) to maximize (1); (iv) after demand is realized, the manufacturer chooses production quantities by solving problem (2) and the two supply chain members obtain their profits. Note that the retailer's reporting decision in (ii) is based on how the manufacturer forms her posterior belief, as formalized in the following definition. The retailer's reporting strategy $\gamma : S \to M$ and the manufacturer's capacity levels (K_1^*, K_2^*) and belief system $\mu(a, b|m)$ constitute a PBE if the following conditions hold:

- For any retailer type $(a, b) \in S$, $\gamma(a, b) \in \arg\max_{m\in M} \Pi_r(K_1^*, K_2^*)$;
- for any message $m \in M$, $(K_1^*, K_2^*) \in \arg\max_{K_1,K_2\geq 0} \Pi_m(K_1, K_2|m)$ as defined in (1);
- the manufacturer's belief is updated per Bayes' rule on the equilibrium path, i.e., $\mu(a, b|m) = \frac{I(\gamma(a,b)=m)\tilde{g}(a,b)}{\sum_{b'\in\{b_l,b_h\}} \int_{\underline{a}}^{\bar{a}} I(\gamma(a',b')=m)\tilde{g}(a',b')da'}$ where $I(x)$ is an indicator function which equals one (resp. zero) if statement x is true (resp. false), and $\tilde{g}(a, b)$ represents the prior joint density function of (a, b), i.e., $\tilde{g}(a, b_t) = \rho_t g(a)$ for $t = h, l$. For messages off equilibrium paths, i.e., m' with $I(\gamma(a, b) = m') = 0$ for all $(a, b) \in S$, no requirement is imposed on $\mu(a, b|m')$.

In any cheap talk game, a babbling equilibrium always exists in which the communication is not credible and hence not informative at all such that the receiver's belief remains as the prior. As is common in the cheap talk literature (Chu et al., 2017; Berman et al., 2019; Lu and Tomlin, 2022), one key goal is to show whether and under what conditions an informative equilibrium exists. An informative equilibrium exists if there is a partition $\{S_1, S_2\}$ of the type space S such that the retailer will report $\gamma(a, b) = m$ for $(a, b) \in S_1$ but report $\gamma(a', b') = m'$ for $(a', b') \in S_2$ where $m \neq m'$. Moreover, we say an informative equilibrium is *influential* if the manufacturer chooses different capacity levels in response to different messages m and m'. As we will establish below, an informative and influential equilibrium can exist under quite general conditions.

4 Informative Equilibrium and Implications

4.1 The Manufacturer's Optimal Capacity Levels

Given any message m, the manufacturer's optimal capacity levels K_1^* and K_2^* will depend on her posterior belief $\mu(a,b|m)$ of the demand distribution. In general, we can define the posterior demand distribution as $F_D(x|m) = \sum_{t\in\{l,h\}} \int_{\underline{a}}^{\bar{a}} F((x-a)b_t)\mu(a,b_t|m)da$ and its complement as $\bar{F}_D(x|m) = 1 - F_D(x|m)$. The optimal capacity levels K_1^* and K_2^* are characterized in the following lemma. We note that similar results have been documented in the literature (e.g., Martínez-de Albéniz and Simchi-Levi, 2005; Fu et al., 2010) in which the authors study optimization problems from the manufacturer's perspective. In our paper, the result will serve as a building block for the analysis of the cheap-talk game between the manufacturer and the retailer.

Lemma 1. *Given any message m, the manufacturer's optimal capacity levels K_1^* and K_2^* are characterized as follows.*

(1) If $r_2 + c_2 \geq r_1 + c_1$, then K_1^* solves the equation $\bar{F}_D(K_1^*|m) = \frac{r_1}{w-c_1}$ and $K_2^* = 0$.

(2) If $r_2 + c_2 < r_1 + c_1$ and $\frac{r_1}{w-c_1} \geq \frac{r_2}{w-c_2}$, then $K_1^* = 0$ and K_2^* solves the equation $\bar{F}_D(K_2^*|m) = \frac{r_2}{w-c_2}$.

(3) If $r_2 + c_2 < r_1 + c_1$ and $\frac{r_1}{w-c_1} < \frac{r_2}{w-c_2}$, then K_1^* and K_2^* are determined by the following equations: $\bar{F}_D(K_1^* + K_2^*|m) = \frac{r_1}{w-c_1}$ and $\bar{F}_D(K_2^*|m) = \frac{r_2-r_1}{c_1-c_2}$.

Depending on the reservation and execution costs of the capacity sources, the manufacturer procures capacity from only one source or from both sources. Note that if the manufacturer procures only source-i capacity then the newsvendor critical fractile is $1 - \frac{r_i}{w-c_i}$. Recall that by definition source 1 (the more flexible source) has the lower reservation cost, i.e., $r_1 < r_2$. In Case (1) of Lemma 1, source 1 is also cheaper in its total cost, i.e., reservation plus execution. Therefore, source 1 dominates source 2 – and is the only procured source – because the profit of any capacity portfolio that procures both sources can be improved upon by trading a unit of source 2 capacity for a unit of source 1 capacity. In Case (2) of Lemma 1, source 2 is cheaper in total cost and also has a higher critical fractile (resulting in a higher service level). It therefore dominates source 1 because the profit of any capacity portfolio that procures both sources can be improved upon by trading a unit of source 1 capacity for a unit of source 2 capacity. In Case (3), neither source dominates and the manufacturer procures capacity from both sources: taking advantage of source 2's lower total cost but also source 1's lower reservation cost and higher critical fractile to enable a higher service level. In this case, source 1's critical fractile determines the total portfolio capacity.

4.2 Strategic Communication of the Retailer's Forecast (a, b)

The retailer, with private forecast information consisting of both the average demand a and the forecast accuracy b, decides on message m to send, anticipating that the manufacturer will optimize capacity levels according to Lemma 1 given her posterior belief $\mu(a, b|m)$.

4.2.1 Existence of Informative Equilibrium The entire space of the retailer's type can be represented by two line segments as shown in Figure 1. Each line segment corresponds to an accuracy level (b_h or b_l) while the continuum of points on each line segment represents the range of average demand $[\underline{a}, \bar{a}]$. Given any two points $x_h \in [\underline{a}, \bar{a}]$ and $x_l \in [\underline{a}, \bar{a}]$ on line segments corresponding to b_h and b_l, respectively, the type space S can be partitioned into two subspaces defined as: $S_1(x_l, x_h) = \{(a, b) \in S|a \leq x_h I(b = b_h) + x_l I(b = b_l)\}$ and $S_2(x_l, x_h) = \{(a, b) \in S|a > x_h I(b = b_h) + x_l I(b = b_l)\}$, where $I(\cdot)$ is an indicator function. As such, $S_1(x_l, x_h)$ and $S_2(x_l, x_h)$ denote the subspaces on the $\underline{a}$ end and the $\bar{a}$ end (i.e., on the left and right in Figure 1), respectively.[3]

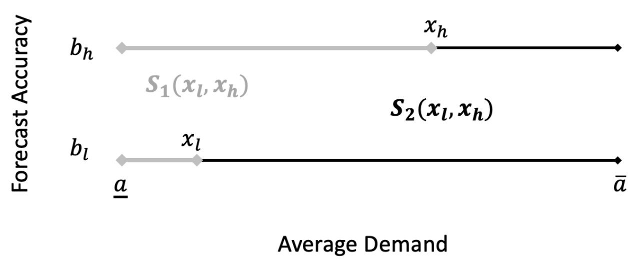

Figure 1: Illustration of type partition: $S_1(x_l, x_h) = \{(a, b) \in S|a \leq x_h I(b = b_h) + x_l I(b = b_l)\}$ in grey; $S_2(x_l, x_h) = \{(a, b) \in S|a > x_h I(b = b_h) + x_l I(b = b_l)\}$ in black.

We focus on a class of message spaces in which, given some (x_l, x_h), the retailer self-reports to which subspace his private forecast (a, b) belongs. That is, we consider the message space as $M = \{S_i(x_l, x_h)|i \in \{1, 2\}, x_l, x_h \in [\underline{a}, \bar{a}]\}$. For example, by sending a message $m = S_1(x_l, x_h)$, the retailer indicates "my average demand a and forecast accuracy b fall into set $S_1(x_l, x_h)$." We are interested in whether there are some $x_l, x_h \in [\underline{a}, \bar{a}]$ such that an informative equilibrium exists

[3]By parameterizing partitions with a pair of (x_l, x_h), we have restricted attention to contiguous two-region partitions in the sense that within each subspace, the set of a is convex on each accuracy level. Theoretically, there can be noncontiguous partitions involving a nonconvex set of a on some accuracy level; see an example in Appendix I. However, such noncontiguous partitions may lack managerial interpretation and therefore be difficult to implement in practice. We therefore focus on contiguous two-region partitions throughout the paper.

in which the retailer will truthfully report $m = S_i(x_l, x_h)$ if his forecast $(a, b) \in S_i(x_l, x_h)$ where $i = 1, 2$.

In any informative equilibrium in which the retailer reports $m = S_1(x_l, x_h)$, the resulting posterior belief of the manufacturer is given by

$$\mu(a, b_t | S_1(x_l, x_h)) = \frac{\rho_t G(x_t)}{\rho_l G(x_l) + \rho_h G(x_h)} \cdot \frac{g(a)}{G(x_t)} = \frac{\rho_t g(a)}{\rho_l G(x_l) + \rho_h G(x_h)}, \ \forall t \in \{l, h\}, a \in [\underline{a}, x_t].$$

Note that for any $t \in \{l, h\}$, $\frac{\rho_t G(x_t)}{\rho_l G(x_l)+\rho_h G(x_h)}$ is the conditional probability of $b = b_t$ given $m = S_1(x_l, x_h)$, and $\frac{g(a)}{G(x_t)}$ is the conditional probability density of a given $b = b_t$ and m, i.e., the original density truncated by interval $[\underline{a}, x_t]$.[4] Similarly, for $m = S_2(x_l, x_h)$ we have

$$\mu(a, b_t | S_2(x_l, x_h)) = \frac{\rho_t \bar{G}(x_t)}{\rho_l \bar{G}(x_l) + \rho_h \bar{G}(x_h)} \cdot \frac{g(a)}{\bar{G}(x_t)} = \frac{\rho_t g(a)}{\rho_l \bar{G}(x_l) + \rho_h \bar{G}(x_h)}, \ \forall t \in \{l, h\}, a \in [x_t, \bar{a}],$$

where $\bar{G}(x) = 1 - G(x)$. The retailer has no incentive to lie if under the two posterior beliefs $\mu(a, b_t|S_1(x_l, x_h))$ and $\mu(a, b_t|S_2(x_l, x_h))$, the manufacturer will choose the same level of total capacity $K_T = K_1 + K_2$, because the retailer's expected payoff (3) depends on the manufacturer's capacity levels only through K_T. In other words, for any informative equilibrium to exist, the manufacturer's total capacity level must be message-independent. Otherwise, the retailer will choose the message leading to the highest total capacity, regardless of the actual forecast. Furthermore, we can show that this message-independent total capacity level is also optimal for the manufacturer in the case of no communication (such that the retailer will not be worse off engaging in communication).

Lemma 2. *In any informative equilibrium, the manufacturer's total optimal capacity level is message-independent, i.e.,* $K_T^*(S_1) = K_T^*(S_2)$.[5] *Furthermore,* $K_T^*(S_1) = K_T^*(S_2)$ *are also optimal for the manufacturer's problem without communication.*

In the following, we will show that an informative equilibrium always exists, and moreover, although informative communication does not change the total capacity, it can be *influential* because different messages enable the manufacturer to choose different capacity combinations, i.e., $K_i^*(S_1) \neq K_i^*(S_2)$ for both $i = 1, 2$.

Consider the case where neither capacity source dominates, i.e., case (3) of Lemma 1. As we established in Lemma 1, the manufacturer's optimal total capacity is determined by a single equation

[4] By the original definition, the conditional density given $b = b_t$ and S_1 is $\frac{\rho_t g(a)}{\rho_t G(x_t)} = \frac{g(a)}{G(x_t)}$ where probability ρ_t is canceled.

[5] For ease of notation, we suppress the dependence of S_i on (x_l, x_h) in the manufacturer's optimal capacity levels $K_i^*(S_j(x_l, x_h))$.

$\bar{F}_D(K_T^*|m) = \frac{r_1}{w-c_1}$. We can express $\bar{F}_D(k|m)$ given $m = S_i(x_l, x_h)$ as a function $\Psi_i(k; x_l, x_h)$ defined below.

$$\begin{aligned}\Psi_1(k; x_l, x_h) &= \int_{\underline{a}}^{x_l} \bar{F}((k-a)b_l)\frac{\rho_l g(a)}{\rho_l G(x_l)+\rho_h G(x_h)}da + \int_{\underline{a}}^{x_h} \bar{F}((k-a)b_h)\frac{\rho_h g(a)}{\rho_l G(x_l)+\rho_h G(x_h)}da,\\ \Psi_2(k; x_l, x_h) &= \int_{x_l}^{\bar{a}} \bar{F}((k-a)b_l)\frac{\rho_l g(a)}{\rho_l \bar{G}(x_l)+\rho_h \bar{G}(x_h)}da + \int_{x_h}^{\bar{a}} \bar{F}((k-a)b_h)\frac{\rho_h g(a)}{\rho_l \bar{G}(x_l)+\rho_h \bar{G}(x_h)}da.\end{aligned} \tag{4}$$

As a result, an informative equilibrium exists if one can find some $x_l, x_h \in [\underline{a}, \bar{a}]$ and K_T which solve the following system of equations:

$$\frac{r_1}{w-c_1} = \Psi_1(K_T; x_l, x_h) \text{ and } \frac{r_1}{w-c_1} = \Psi_2(K_T; x_l, x_h). \tag{5}$$

The two equations in (5) correspond to the manufacturer's optimal total capacity in response to $m = S_1(x_l, x_h)$ and $m = S_2(x_l, x_h)$, respectively. In fact, the existence of a pair (x_l, x_h) satisfying (5) is guaranteed as long as a is a continuous random variable as we have assumed. This can be proved by applying the Intermediate Value Theorem. An analogous result holds for the cases where only one capacity source is used, i.e., cases (1) and (2) of Lemma 1.

Defining $a_m = \frac{\underline{a}+\bar{a}}{2}$ as the mid-point of $[\underline{a}, \bar{a}]$, a key result is stated in the following proposition.

Proposition 1. *There exists $y \in [-\frac{\bar{a}-\underline{a}}{2}, \frac{\bar{a}-\underline{a}}{2}]$ such that an informative equilibrium exists in which for $i = 1, 2$, the retailer with private forecast $(a, b) \in S_i(x_l, x_h)$ reports $m = S_i(x_l, x_h)$ where $x_l = a_m - y$ and $x_h = a_m + y$.*

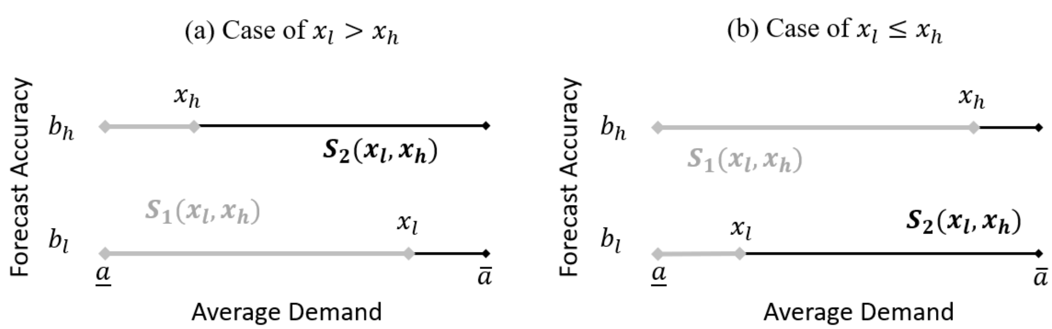

Figure 2: Graphic illustration of informative partitions

Proposition 1 establishes that, in general, the retailer may in equilibrium truthfully report the subspace to which his forecast belongs. Moreover, one can always construct a partition in a symmetric manner (i.e., with $x_l + x_h = 2a_m$) to support an informative equilibrium. Throughout this section, we will focus on symmetric partitions as characterized in Proposition 1. As discussed later in §5.3, there can be asymmetric partitions that support informative equilibra, but our main findings based on symmetric partitions remain to hold under asymmetric partitions.

The partition characterized in Proposition 1 takes on one of two possible forms depending on whether $y < 0$ or $y \geq 0$: when $y < 0$ we have $x_l > x_h$ but we have $x_l \leq x_h$ when $y \geq 0$. The two possible cases are illustrated in Figure 2. In either case, from the manufacturer's perspective, the posterior expected average demand conditional on receiving message $S_1(x_l, x_h)$ is lower than that conditional on receiving $S_2(x_l, x_h)$. Consider the case of $x_l > x_h$. Comparing the messages $S_1(x_l, x_h)$ and $S_2(x_l, x_h)$, the message $S_1(x_l, x_h)$ consists of a shorter interval of a (i.e., $[\underline{a}, x_h]$) at the high accuracy b_h but a longer interval of a (i.e., $[\underline{a}, x_l]$) at the low accuracy b_l. Therefore, the expected accuracy conditional on $S_1(x_l, x_h)$ is lower than the expected accuracy conditional on $S_2(x_l, x_h)$. Therefore, when $x_l > x_h$, a message $m = S_1(x_l, x_h)$ can be viewed as a statement implying that "in your expectation I have a lower average demand and a lower forecast accuracy" than implied by a message $m = S_2(x_l, x_h)$. For expositional brevity, when $x_l > x_h$, we refer to $S_1(x_l, x_h)$ as a "low-average, low-accuracy" message and refer to $S_2(x_l, x_h)$ as a "high-average, high-accuracy" message. Analogously, and again based on conditional expectations, when $x_l \leq x_h$, a message $m = S_1(x_l, x_h)$ represents a "low-average, high-accuracy" message whereas $m = S_2(x_l, x_h)$ represents a "high-average, low-accuracy" message.

4.2.2 When the Informative Equilibrium Is also Influential Next, we will elaborate on why these forms of messages can result in informative communication and show that such communication can be influential. For the moment, let us consider a special case where the prior distribution of a is uniform over $[\underline{a}, \bar{a}]$ and the random noise ϵ also follows a uniform distribution with support $[-\sigma, \sigma]$. We define $\phi = \frac{r_1}{w - c_1}$ and

$$y^o = \begin{cases} \frac{4\sigma\rho_h\rho_l(b_h - b_l)(1-2\phi) - \sqrt{(\bar{a}-\underline{a})^2(\rho_l b_l + \rho_h b_h)^4 + 16\sigma^2\rho_h^2\rho_l^2(b_h-b_l)^2(1-2\phi)^2}}{2(\rho_l b_l + \rho_h b_h)^2} & \text{for } \phi \leq \frac{1}{2} \\ \frac{4\sigma\rho_h\rho_l(b_h - b_l)(1-2\phi) + \sqrt{(\bar{a}-\underline{a})^2(\rho_l b_l + \rho_h b_h)^4 + 16\sigma^2\rho_h^2\rho_l^2(b_h-b_l)^2(1-2\phi)^2}}{2(\rho_l b_l + \rho_h b_h)^2} & \text{for } \phi > \frac{1}{2}. \end{cases} \tag{6}$$

Additionally, define $\Psi_0(k) = \sum_{t=l,h} \rho_t \int_{\underline{a}}^{\bar{a}} \bar{F}((k-a)b_t)dG(a)$. Note that $\Psi_0(k)$ represents the shortage probability given a total capacity level k under the prior demand distribution, whereas Ψ_1 and Ψ_2 defined in (4) represent the shortage probabilities conditioning on S_1 and S_2, respectively. We impose two regularity conditions: (C1) $\Psi_0(\underline{a} + \frac{\sigma}{b_h}) \leq \phi$; (C2) $\min_{i=1,2} \Psi_i(\bar{a} - \frac{\sigma}{b_h}; a_m - y^o, a_m + y^o) \geq \frac{r_2 - r_1}{c_1 - c_2}$. These conditions ensure that the equilibrium capacity levels K_T^* and K_2^* will not cover or leave out all the demand realizations, i.e., for all $(a, b) \in S$, $(K_T^* - a)b$ and $(K_2^*(m) - a)b$ are within

$[-\sigma, \sigma]$. In Appendix D, we show that for the case of $\rho_h = \rho_l = \frac{1}{2}$, these conditions hold if the support of ϵ is sufficiently wide compared to the support of a, i.e., $\bar{a} - \underline{a}$ is small enough relative to 2σ. Note that these regularity conditions are assumed only for analytical convenience. Without these conditions, one can always numerically determine the values of y and (x_l, x_h) that support an informative equilibrium per Proposition 1.

Proposition 2. *Consider the case in which G is a uniform distribution over $[\underline{a}, \bar{a}]$ and F is a uniform distribution over $[-\sigma, \sigma]$. (i) if neither capacity source dominates the other (i.e., $r_2 + c_2 < r_1 + c_1$ and $\frac{r_1}{w-c_1} < \frac{r_2}{w-c_2}$) and regularity conditions (C1)(C2) hold, then an informative equilibrium exists with $x_l = a_m - y^o$ and $x_h = a_m + y^o$ in which the retailer reports $m = S_i(x_l, x_h)$ if his private forecast $(a, b) \in S_i(x_l, x_h)$ for all $i = 1, 2$; moreover, the informative equilibrium is influential as characterized below:*

(a) If $\phi \leq \frac{1}{2}$, then we have $x_l > x_h$ and the manufacturer's optimal capacity levels satisfy $K_1^*(S_1) > K_1^*(S_2)$ and $K_2^*(S_1) < K_2^*(S_2)$.

(b) If $\phi > \frac{1}{2}$, then we have $x_l < x_h$ and the manufacturer's optimal capacity levels satisfy $K_1^*(S_1) < K_1^*(S_2)$ and $K_2^*(S_1) > K_2^*(S_2)$.

(ii) If one of the capacity sources dominates the other (i.e., $r_2 + c_2 \geq r_1 + c_1$ or $\frac{r_1}{w-c_1} \geq \frac{r_2}{w-c_2}$), then an informative but noninfluential equilibrium exists.

Proposition 2 provides a closed-form characterization of the informative partition. Importantly, part (i) establishes that when neither capacity source dominates, informative communication is *influential*: the manufacturer will choose different capacity levels in response to different retailer messages. Note that under the optimal capacity portfolio, $\phi = \frac{r_1}{w-c_1}$ equals the shortage probability (when neither source dominates) and $1 - \frac{r_1}{w-c_1}$, the newsvendor fractile, is the optimal service level.

First, consider case (a) of part (i), that is, $\phi \leq \frac{1}{2}$. We have $x_l > x_h$, and therefore, as discussed after Proposition 1, $S_1(x_l, x_h)$ represents a "low-average, low-accuracy" retailer message and $S_2(x_l, x_h)$ represents a "high-average, high-accuracy" message. When $\phi \leq \frac{1}{2}$, the optimal service level is greater than or equal to 50% and thus the optimal total capacity level increases in demand variability. In this case, intuitively, the two messages, "low-average, low-accuracy" and "high-average, high-accuracy," can result in the same total capacity level such that the retailer has no incentive to lie. However, when choosing how to reserve the same total capacity, it is valuable to the manufacturer to know in which subspace the retailer's forecast resides. It will choose to reserve more of the flexible capacity (source 1) when receiving the "low-average, low-accuracy" message S_1 than when receiving the "high-average, high-accuracy" message S_2. Likewise, it will reserve more of the inflexible capacity (source 2) under message S_2 than under message S_1. That is, $K_1^*(S_1) > K_1^*(S_2)$ and $K_2^*(S_1) < K_2^*(S_2)$.

Next consider case (b) of part (i), i.e., $\phi > \frac{1}{2}$. We have $x_l < x_h$ and, therefore, as discussed after Proposition 1, $S_1(x_l, x_h)$ represents a "low-average, high-accuracy" message and $S_2(x_l, x_h)$ represents a "high-average, low-accuracy" message. When $\phi > \frac{1}{2}$, the optimal service level is less than 50% and thus the optimal total capacity level decreases in demand variability. In this case, the same total capacity level can result from either a "low-average, high-accuracy" message or a "high-average, low-accuracy" message. As before, the retailer has no incentive to lie but again the manufacturer will react differently to the two messages: reserving less flexible capacity under the "low-average, high-accuracy" message than under the "high-average, low-accuracy" message and vice versa for the inflexible capacity. That is, $K_1^*(S_1) < K_1^*(S_2)$ and $K_2^*(S_1) > K_2^*(S_2)$.

Corollary 1. *Under the informative equilibrium characterized in Proposition 2, as ϕ becomes closer to $\frac{1}{2}$, $|x_h - x_l|$ increases; moreover, at $\phi = \frac{1}{2}$, $|x_h - x_l|$ attains the maximum value $\bar{a} - \underline{a}$ with $x_l = \bar{a}$ and $x_h = \underline{a}$, whereas the minimum value of $|x_h - x_l|$ is nonzero and attained as $\phi \to 0$ or 1.*

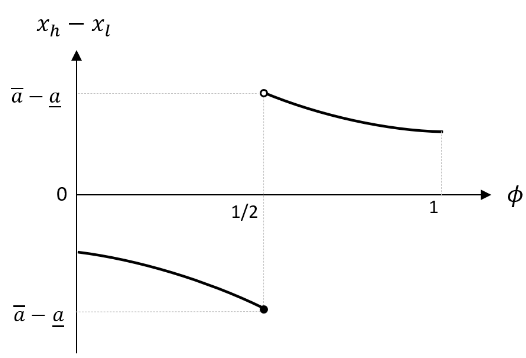

Figure 3: Illustration of $x_h - x_l$ as a function of ϕ.

Corollary 1, as illustrated in Figure 3, shows that as ϕ becomes closer to 1/2, x_h and x_l move further away from each other. In particular, when $\phi = \frac{1}{2}$, $(x_l, x_h) = (\bar{a}, \underline{a})$ such that the equilibrium messages reduce to simply reporting whether $b = b_l$ or $b = b_h$, i.e., forecast accuracy b is fully revealed.[6] To provide an intuitive explanation for this, recall that in a newsvendor problem with a critical fractile $\frac{1}{2}$ (and a symmetric demand distribution), the optimal capacity depends only

[6]Note that in the special case of $\phi = \frac{1}{2}$, both $(x_l, x_h) = (\bar{a}, \underline{a})$ and $(x_l, x_h) = (\underline{a}, \bar{a})$ can constitute an informative partition such that b is fully revealed. We break the tie by choosing $(x_l, x_h) = (\bar{a}, \underline{a})$. This is without loss of generality, since setting $(x_l, x_h) = (\underline{a}, \bar{a})$ instead is equivalent to swapping the current indexes of S_1 and S_2.

on the average demand, independent of the variance. In our context, the manufacturer solves a similar newsvendor problem while choosing the total capacity level based on a critical fractile $1-\phi$. Therefore, in the special case of $\phi = \frac{1}{2}$, the retailer is willing to fully reveal his forecast accuracy, because the accuracy has no impact on the manufacturer's total capacity. Moreover, as ϕ deviates from 0.5 in either direction, the difference between x_l and x_h shrinks and so less forecast-accuracy information will be revealed.[7] As ϕ goes to 0 or 1, x_l and x_h will have the minimum distance but cannot coincide, since a partition with $x_l = x_h$ reveals only the demand-average information and cannot be incentive-compatible.[8]

4.2.3 Value of Communication We now return to the general case. Importantly, and as formally established in the following proposition (subject to a very mild restriction on the forecast noise distribution), influential communication strictly improves the manufacturer's expected profit as compared to the case without any communication. The reason is that communication enables the manufacturer to tailor her capacity portfolio to the subspace in which the demand forecast (a, b) resides instead of using a one-size-fits-all capacity configuration based only on the prior knowledge of (a, b). The retailer's expected payoff, on the other hand, is not impacted by this influential communication because (as shown in Lemma 2) the manufacturer's total capacity level with influential communication is the same as in the case without any communication.

Proposition 3. *Assume that F is strictly increasing over its support. Compared to the case without any communication, influential communication, whenever it occurs in equilibrium, will make the manufacturer strictly better off while maintaining the retailer's payoff unchanged. Therefore, the supply chain members' payoffs are Pareto-improved by cheap talk communication.*

The assumption of a strictly increasing F ensures that the manufacturer's capacity reservation problem always has a unique solution so that the manufacturer is strictly better off with message-dependent capacity configurations. Cheap talk (which incurs no cost and requires no investment) can lead to a Pareto-improvement of the supply chain payoffs. Proposition 3 also provides support for informative communication to be a reasonable outcome despite the multiplicity of equilibria in cheap talk games. It has been shown that in any cheap talk game a babbling equilibrium, in

[7] The amount of forecast-accuracy information being revealed can be measured by the variance of b conditional on messages. The variance of a two-point distributed b is equal to $v(1-v)(b_h - b_l)^2$ where v represents the conditional probability of $b = b_l$, that is, $v = \frac{\rho_l G(x_l)}{\rho_l G(x_l) + \rho_h G(x_h)}$ conditional on $S_1(x_l, x_h)$ whereas $v = \frac{\rho_l \bar{G}(x_l)}{\rho_l \bar{G}(x_l) + \rho_h \bar{G}(x_h)}$ conditional on $S_2(x_l, x_h)$. The variance is zero when x_l and x_h are respectively at the two end points of $[\underline{a}, \bar{a}]$ (i.e., $v = 1$ or $v = 0$), and increases as x_l and x_h move toward the midpoint of $[\underline{a}, \bar{a}]$.

[8] If $x_l = x_h$, then $S_1 = \{(a, b) \in S | a \leq a_m\}$ and $S_2 = \{(a, b) \in S | a > a_m\}$. As a result, S_2 will imply a higher demand average with the same accuracy information compared to S_1, so the retailer always prefers to report $m = S_2$.

which communication is not informative at all, always exists per the definition of PBE (Sobel, 2020). However, as proved, the informative equilibrium results in a Pareto-improvement and is thus preferable to the babbling equilibrium based on the Pareto-dominance rule commonly used in equilibrium selection.

Table 1: Numerical Examples of Influential Communication [Parameter values: $\epsilon \sim U[-500, 500]$, $a \sim U[500, 600]$, $(r_1, c_1) = (1, 4.5)$, $(r_2, c_2) = (5, 0)$, $b_l = 1, b_h = 10$ and $\rho_l = \rho_h = 0.5$]

		No Communication		**With Communication**			
w	ϕ	K_1^N, K_2^N	Π_M^N	$K_1^*(S_1), K_2^*(S_1)$	$K_1^*(S_2), K_2^*(S_2)$	Π_M^*	$\frac{\Pi_m^* - \Pi_m^N}{\Pi_m^N}$ (%)
7	0.40	297.8, 272.2	729.4	396.5, 173.5	73.3, 496.7	771.1	5.71
6.5	0.50	277.8, 272.2	487.5	388.9, 161.1	52.9, 497.1	553.7	13.58
6	0.67	241.5, 272.2	252.3	74.0, 439.7	300.5, 213.2	263.4	4.40

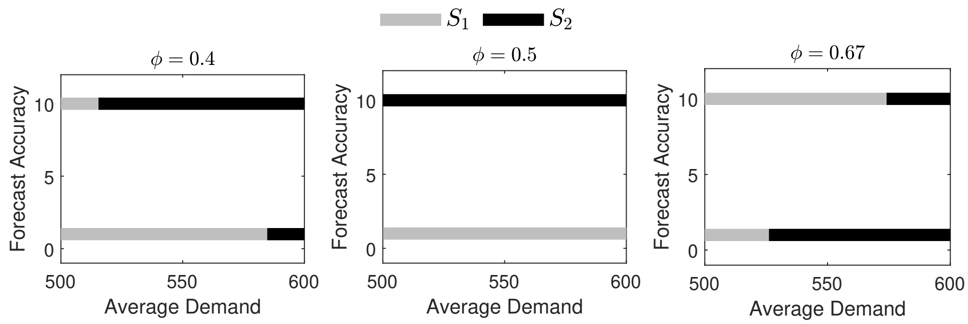

Figure 4: Informative partitions of the numerical examples in Table 1

We illustrate our findings and explore the profit implications with a numerical example presented in Table 1. We consider three instances with different wholesale prices; $w = 6$, 6.5 and 7. The parameters are chosen such that neither capacity source dominates, and so the manufacturer will procure both capacity sources in the optimal portfolio. For the three different values of w, $\phi = \frac{r_1}{w - c_1}$ is given by 0.4, 0.5, and 0.67, respectively. We compute the informative partitions for each of the three instances; see Figure 4. The partitions echo our analytical results. When $\phi = 0.4$, credible communication arises in equilibrium with a "low-average, low-accuracy" versus "high-average, high-accuracy" partition as in case (i-a) of Proposition 2, whereas when $\phi = 0.67$, credible communication is supported by a "low-average, high-accuracy" versus "high-average, low-accuracy" partition as in case (i-b) of Proposition 2. When $\phi = 0.5$, the equilibrium messages reduce to simply reporting the true accuracy level b.

We also numerically compute and report the manufacturer's capacity levels (K_1^N, K_2^N) and expected profits Π_M^N without any communication, i.e., based on the prior belief of the retailer's forecast

(a, b) (see the No Communication columns of Table 1), and the manufacturer's message-dependent capacity levels $(K_1^*(m), K_2^*(m))$ and expected profits Π_M^* in the informative and influential equilibrium (see the With Communication columns of Table 1). As reported in Table 1, the manufacturer chooses very different capacity portfolios $(K_1^*(m), K_2^*(m))$ under messages $m = S_1$ and $m = S_2$, which showcases that cheap talk communication can substantially influence the manufacturer's capacity decisions. (Also, as established theoretically, the total capacity $K_1^*(m) + K_2^*(m)$ is constant for different messages and so the retailer has no incentive to lie.) For example, in the case of $\phi = 0.4$, the manufacturer, if receiving a low-average, low-accuracy message S_1, will reserve more capacity from the flexible source, i.e., $K_1^*(S_1) = 396.5 > K_2^*(S_1) = 173.5$; however, if receiving a high-average, high-accuracy message S_2, she will do the opposite $K_1^*(S_2) = 73.3 < K_2^*(S_2) = 496.7$. These message-dependent capacity combinations also differ significantly from (K_1^N, K_2^N), the manufacturer's optimal capacity levels absent of any communication.

Furthermore, comparing Π_M^N and Π_M^* in Table 1, we can see that the manufacturer obtains a higher expected profit with influential communication than without communication. We also observe that the percentage improvement of the manufacturer's payoff is the highest at $w = 6.5$, i.e., when $\phi = 0.5$. This is because, as proved in Corollary 1, the retailer's forecast accuracy b is fully revealed in equilibrium when $\phi = 0.5$. Consequently, the manufacturer can configure her capacity portfolio based on full information about the forecast accuracy, and thus receive the greatest gain from communication. Our observations imply that the value of communication stems from the information revelation of forecast accuracy, not the average demand. The reason is that in equilibrium the manufacturer's total capacity level is not responsive to different messages and so the updated average-demand information is not leveraged. (However, the dimension of private average-demand information is essential for an informative equilibrium to exist.) We note that although the magnitude of the improvement in the manufacturer's payoff depends on parameter values, the improvement can be achieved at no cost because it is the outcome of pure cheap talk. Therefore, our results indicate that money is being left on the table if upstream firms are unaware of the value of eliciting a downstream firm's demand forecast (that includes both the average demand and the forecast accuracy) through free and nonbinding communication.

A more extensive numerical study is reported in Appendix E.1 where we consider the case in which ϵ follows a truncated normal distribution. This additional numerical study confirms that the above observations on the value of communication are robust. It is worth noting that accuracy information is fully revealed specifically at $\phi = 0.5$ because ϵ is assumed to follow symmetric distributions. In Appendix E.2, we numerically show that with asymmetrically distributed ϵ, the value of communication is still maximized when accuracy information is fully revealed, but this happens at some $\phi > 0.5$ (resp. $\phi < 0.5$) when the distribution of ϵ is left-skewed (resp. right-skewed).

4.2.4 Remarks

In closing this subsection, we make the following remarks.

Remark 1. An important feature of our model is that the retailer's profit depends on the manufacturer's capacity portfolio *only* through the total capacity level. This feature is essential for an informative equilibrium to exist, because the retailer has no incentive to lie between the two partitions inducing the same total capacity. As discussed in Appendix H, although we focus on a simple wholesale price contract, our results apply to other given contract agreements that retain this feature such as revenue-sharing contracts and under-delivery penalty contracts.

Remark 2. We focus on an exogenous wholesale price w and our results are established for any given wholesale price w. Hence, if w is endogenously determined before the retailer observes and communicates his private forecast, then our main findings continue to hold. Alternatively, w can be determined by the manufacturer after the retailer reports his private forecast. Chu et al. (2017) have analyzed such a setting when the retailer possesses private information only about average demand. They find that because in this case the manufacturer can choose w in response to the retailer's report, cheap-talk communication can be truthful under certain conditions. It is a valuable question for future research whether including forecast accuracy as an additional dimension of private information will enhance the communication credibility when w is determined after communication.

Remark 3. It is worth noting that the message m in our model need not be the exact set S_1 or S_2; it can be plain words or simpler measures as long as the manufacturer can map each message to the corresponding set of types given the common knowledge of model parameters.[9] For example, analogous to Amazon's practice of sharing forecast quantiles (Intentwise, Inc., n.d.), a retailer with type $(a, b) \in S_i$ could report a number of quantiles based on the conditional demand distribution $E[D|S_i]$. Such communication will be credible, provided that the manufacturer can map the quantile information to the corresponding set S_i.

[9] By the definition of PBE, the manufacturer needs to form a consistent belief $\mu(a, b|m)$ given each message m in equilibrium where message m can take any form. For example, in the case of $x_l > x_h$, $S_1(x_l, x_h)$ and $S_2(x_l, x_h)$ represent "low-average, low-accuracy" and "high-average, high-accuracy" subspaces, respectively. The message space can be defined simply as $M = \{$"low-average, low-accuracy", "high-average, high-accuracy"$\}$. Under a belief system such that the manufacturer anticipates "low-average, low-accuracy" sent from a type $(a, b) \in S_1(x_l, x_h)$ and "high-average, high-accuracy" sent from a type $(a, b) \in S_2(x_l, x_h)$, the simple message space M will support a PBE which is essentially equivalent to that with $M = \{S_1(x_l, x_h), S_2(x_l, x_h)\}$.

4.3 Effect of Capacity Costs

Intuitively one might anticipate that the manufacturer would always be worse off with a higher reservation or execution cost. However, such intuition is not necessarily correct when the value of communication is considered. On the one hand, an increase in these costs has a direct negative impact because the manufacturer pays more for the capacity sources. On the other hand, as discussed above, a change in these costs may influence the value of ϕ, resulting in more information revelation (with respect to the forecast accuracy) if ϕ becomes closer to $\frac{1}{2}$. In the following proposition, we show that, perhaps surprisingly, the manufacturer can be better off if the flexible source (i.e., source 1) becomes more expensive to reserve or execute.

Proposition 4. *Suppose that both capacity sources are used (i.e., $r_1+c_1 > r_2+c_2$ and $\frac{r_1}{w-c_1} < \frac{r_2}{w-c_2}$) and $\rho_h = \rho_l = \frac{1}{2}$. The manufacturer's expected profit is always decreasing in r_2 and c_2, but can be increasing in r_1 and c_1. In particular, given $a \sim U[\underline{a}, \bar{a}]$, $\epsilon \sim U[-\sigma, \sigma]$ and the regularity conditions of Proposition 2(i), there exists a threshold Γ, which is increasing in σ, such that if $\bar{a} - \underline{a} < \Gamma$, the manufacturer's expected profit is increasing in r_1 and c_1 within an interval $\phi = \frac{r_1}{w-c_1} \in [\frac{1}{2} - \delta, \frac{1}{2}]$ for some $\delta > 0$.*[10]

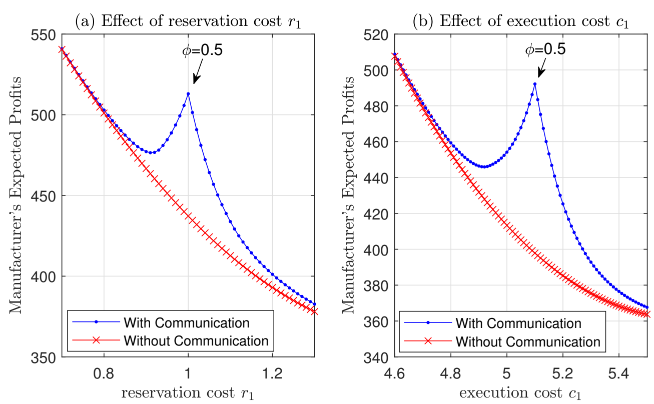

Figure 5: The manufacturer's expected profits as functions of r_1 and c_1. [Parameter values: $a \sim U[500, 510]$, $\rho_h = \rho_l = 1/2$, $b_h = 20$, $b_l = 1$, $\epsilon \sim$ Truncated Normal$(0, 1000^2)$ over $[-500, 500]$, $(r_2, c_2) = (5, 0)$; in panel (a), $c_1 = 4.5$ and r_1 is varied; in panel (b), $r_1 = 0.7$ while c_1 is varied.]

Because ϕ is solely determined by the flexible capacity source, any cost increase associated with the inflexible source (i.e., increase in r_2 or c_2) can only hurt the manufacturer. However an increase in r_1 or c_1 can bring $\phi = \frac{r_1}{w-c_1}$ closer to $\frac{1}{2}$, thereby improving the information revelation during

[10]The closed-form expression of Γ can be found in the proof of Proposition 4. A more general version of Proposition 4 is presented in Proposition A.1 of Appendix A where the threshold is generalized to the case with more than two capacity sources.

communication. Under the special case considered in Proposition 2, we can prove that if the range of the average demand $\bar{a} - \underline{a}$ is small compared to the support of the noise term such that the accuracy information is relatively important, then the informational advantage from an increase in r_1 or c_1 can outweigh the direct cost disadvantage. Figure 5 represents a numerical example with ϵ following a truncated normal distribution, which illustrates the analytical result in Proposition 4. As expected, without communication, the manufacturer is always worse off as r_1 or c_1 increases. However, the manufacturer's expected profit with communication is U-shaped in r_1 for $r_1 \leq 1$ in Panel (a) and in c_1 for $c_1 \leq 5.1$ in Panel (b). The gap between the profits with and without communication represents the value of communication which is maximized when the values of r_1 and c_1 lead to $\phi = 0.5$. In other words, our results imply that a reduction in the cost of the flexible capacity (for example, a reduction in r_1 from 1 to 0.9 in Panel (a)) can hurt the manufacturer if the operational benefit from the cost reduction cannot offset the loss resulting from information revelation.

Additionally, in Appendix G, we present an analysis on the impact of ρ_h, the probability of forecast accuracy being high. We find that the value of communication is maximized as ρ_h takes a moderate value and that an increase in ρ_h generally benefits the manufacturer but not necessarily the retailer.

5 Extensions

5.1 More Than Two Capacity Sources: Implications for Portfolio Expansion

Our results can readily be extended to the case where the manufacturer has access to n sources where $n > 2$. Let r_i and c_i denote the reservation and execution costs, respectively, for capacity source i where $i = 1, 2, ..., n$. We continue to assume that the retailer knows the manufacturer's set of sources and their cost structures. We note that it is quite common for companies to publicly list the location of their factories and/or their external manufacturing sources.[11] Moreover, companies are at times transparent about the comparative cost structures of their manufacturing base.[12] In the literature on asymmetric demand information, it is often assumed that cost information is

[11]See, for example, page 46 of contract manufacturer's Celestica's 2022 annual report and apparel manufacturer VF Corporation's factory list at https://www.vfc.com/responsibility/governance/factory-list

[12]VF Corporation states on page 4 of their 2023 annual report: "products obtained from contractors in the Western Hemisphere generally have a higher cost than products obtained from contractors in Asia. ... The use of contracted production with different geographic regions and cost structures, provides a flexible approach to product sourcing."

common knowledge (Özer et al., 2011; Chu et al., 2017).[13] Without loss of generality, we assume that $w > r_i + c_i$ for all i such that each capacity source is profitable for the manufacturer, and that $r_1 < r_2 < r_3 < ... < r_n$ and $c_1 > c_2 > ... > c_n$.[14] So, as before, a capacity source with a lower index is more flexible in the sense that it has a lower reservation cost.

The manufacturer can now reserve capacity from more than two sources. But similar to the case of $n = 2$, with a capacity portfolio, the manufacturer can leverage the low costs of inflexible sources but maintain a high service level using flexible sources. In Appendix A, we present the detailed model formulation and an algorithm to determine the manufacturer's optimal capacity levels [15] In the lemma below, we highlight a key property of the manufacturer's optimal capacity portfolio

Lemma 3. *Given any posterior demand distribution $F_D(\cdot|m)$, in the optimal solution, the manufacturer always reserves capacity from source κ where $\kappa = \arg\min_{1\leq i\leq n} \frac{r_i}{w-c_i}$. Moreover, the optimal total capacity $K_T^* = \sum_{i=1}^n K_i^* = \bar{F}_D^{-1}(\frac{r_\kappa}{w-c_\kappa}|m)$.*

That is, under any posterior demand distribution, the manufacturer's total capacity is determined solely by the reservation and execution costs associated with source κ, i.e., the source that is associated with the lowest ratio $\frac{r_i}{w-c_i}$ and thus provides the highest service level if being used in isolation. Let us call κ the critical source. As in our base model, the retailer has no incentive to lie between messages $m = S_1$ and $m = S_2$ if both messages induce the manufacturer to choose the same total capacity level. Therefore, our main results can readily be generalized. Moreover, the informative partition depends only on the the critical source κ.

Proposition 5. *Suppose that the manufacturer has $n \geq 2$ capacity sources. There exists $y \in [-\frac{\bar{a}-\underline{a}}{2}, \frac{\bar{a}-\underline{a}}{2}]$ such that an informative equilibrium exists in which, for each $i = 1, 2$, the retailer with private forecast $(a, b) \in S_i(x_l, x_h)$ reports $m = S_i(x_l, x_h)$ where $x_l = a_m - y$, $x_h = a_m + y$. Moreover, the value of y and the message-independent total capacity level K_T can be found by solving equations: $\frac{r_\kappa}{w-c_\kappa} = \Psi_1(K_T; a_m - y, a_m + y) = \Psi_2(K_T; a_m - y, a_m + y)$ where $\kappa = \arg\min_{1\leq i\leq n} \frac{r_i}{w-c_i}$.*

We now explore the value to the manufacturer of expanding its capacity portfolio, that is,

[13] In fact, the essential assumption we need is that the retailer has sufficient knowledge to infer the manufacturer's total capacity given any posterior demand distribution. As we establish in Lemma 3, the manufacturer's total capacity is determined solely by the lowest ratio $\frac{r_i}{w-c_i}$ among all capacity sources i. Therefore, all we need is to assume is that the retailer knows $\min_i \frac{r_i}{w-c_i}$.

[14] If $w \leq r_i + c_i$ for some i, then we can eliminate source i from consideration and re-index the remaining sources. If $c_i < c_j$ and $r_i = r_j$, or $c_i = c_j$ and $r_i < r_j$, for some i and j, it is easy to see that source j is dominated by source i, and thus j can be eliminated from the set of sources under consideration.

[15] We note that given any posterior demand distribution, the manufacturer in our problem solves a portfolio selection problem similar to those studied in Martínez-de Albéniz and Simchi-Levi (2005) and Fu et al. (2010). In our problem, however, the optimal capacity portfolio will serve as the manufacturer's best response to any retailer message m.

expanding the set of sources at which it can build (reserve) capacity. We examine two problem instances that differ only in the capacity portfolio and in doing so we determine the value of having three sources instead of two.

Example 1 (Three sources)**.** The manufacturer has three capacity sources with costs $(r_1, c_1) = (0.6, 5)$, $(r_2, c_2) = (1, 4.5)$, and $(r_3, c_3) = (5, 0)$. The wholesale price $w = 6.5$. The prior for the average demand a follows a uniform distribution $U[500, 600]$; the two accuracy levels are $b_l = 1$ and $b_h = 10$; the noise term ϵ follows a truncated normal $N(0, 500^2)$ over $[-500, 500]$. Using Algorithm A.1 presented in Appendix A, we can show that $I^* = \{1, 2, 3\}$, that is, capacity is reserved at all three sources. The total capacity is determined by critical source $\kappa = 1$, resulting in a service level of $1 - \frac{r_1}{w-c_1} = 60\%$. Solving for the informative partition, we find that $x_l = 585.3$ and $x_h = 514.7$ and therefore the retailer's forecast accuracy is *partially* revealed in equilibrium. The manufacturer's expected profit is $\Pi_m^* = 559.3$.

Example 2 (Two sources)**.** Everything is identical to Example 1 except that source 1 is not in the manufacturer's capacity portfolio, or equivalently $(r_1, c_1) = (\infty, \infty)$. The optimal capacity portfolio becomes $I^* = \{2, 3\}$ with critical source $\kappa = 2$, leading to a service level $1 - \frac{r_2}{w-c_2} = 50\%$. As discussed in §4, this results in $x_l = 600$ and $x_h = 500$ such that the retailer's forecast accuracy is fully revealed. The manufacturer's expected profit is $\Pi_m^* = 568.9$.

Thus, and perhaps surprisingly, the manufacturer is worse off if she has more sources available to her. What explains this? It is driven by the change in information revelation – from full to partial – when the portfolio is expanded to include source 1. Source 2 is the critical source if only 2 and 3 are available but source 1 is the critical source when all three are available. Source 1 provides a higher service level (0.6) than source 2 (0.5). In our earlier two-source model, we established that if the forecast noise is symmetrically distributed, communication brings greatest value when $\frac{r_\kappa}{w-c_\kappa} = 0.5$, because in such a case the retailer's forecast accuracy can be fully revealed through communication. The higher service level achieved by the addition of source 1 reduces the information revelation and therefore hurts the manufacturer even though source 1 is operationally attractive from a reservation cost and service level perspective. In the following proposition, we formalize and generalize this observation regarding the tradeoff between the operational advantage and potential informational disadvantage of having an expanded capacity portfolio. In doing so, we ignore any upfront search or qualification costs associated with portfolio expansion.

Proposition 6. *Consider a firm with n sources, where $n \geq 2$, that expands its portfolio to $n+1$ sources by adding some new source j. All else remains the same. With communication, (i) if the critical source κ is unchanged by the addition of j then the manufacturer's expected profit is (weakly) higher for the expanded portfolio; (ii) if the critical source in the expanded portfolio is $\kappa = j$ then the manufacturer's expected profit can be strictly lower for the expanded portfolio.*[16]

In Appendix A, we show that, similar to the findings in §4.3, the manufacturer's expected profit can be increasing in r_κ and c_κ when the optimal capacity portfolio consists of more than one sources. Altogether our results indicate that information revelation considerations can have a significant impact on the profit implications of capacity portfolio initiatives (set expansion and cost reduction) that would seem to be obviously beneficial from an operations perspective.

5.2 Continuous Accuracy Level

We now consider the case where the forecast accuracy b, like demand average a, also follows a continuous probability distribution with support $[\underline{b}, \bar{b}]$, instead of being two-point distributed. Let $g(a,b)$ represent the prior joint density function of the retailer's forecast (a,b). The entire type space, $S = \{(a,b) | a \in [\underline{a}, \bar{a}], b \in [\underline{b}, \bar{b}]\}$, can be represented as a rectangle as illustrated in Figure 6. Let $(a_m, b_m) = (\frac{\underline{a}+\bar{a}}{2}, \frac{\underline{b}+\bar{b}}{2})$ be the middle point of the rectangle area. With any straight line passing through (a_m, b_m), we can divide the rectangle area into two subspaces. Let y denote the angle measured from the horizontal line $b = b_m$ to this straight line. By varying y from 0 to π, we can rotate the straight line clockwise to construct various partitions. We use $S_1(y)$ (resp. $S_2(y)$) to represent the subspace to the left (resp. right) of the straight line. Specifically, the two subspaces can be explicitly defined with the tangent (tan) and cotangent (cot) functions:

$$S_1(y) = \begin{cases} \{(a,b) \in S | (a_m - a)\tan y \geq b - b_m\} & \text{if } 0 \leq y \leq y^\dagger, \\ \{(a,b) \in S | a_m - a \geq (b - b_m)\cot y\} & \text{if } y^\dagger < y < \pi - y^\dagger, \\ \{(a,b) \in S | (a_m - a)\tan y \leq b - b_m\} & \text{if } \pi - y^\dagger \leq y \leq \pi \end{cases} \tag{7}$$

[16] To be precise, we compare the manufacturer's expected profit in two different problem instances in which one instance has an additional capacity source but otherwise the instances are identical. For expositional ease, we refer to this comparison as capacity expansion. Also, as noted in the statement, adding a source can have a potentially negative impact only if the existing portfolio has at least two sources. The reason is as follows. Recall from §4 that communication is never influential in a portfolio with only one source, and therefore communication does not create value. Thus, adding a second source not only provides a potential operational cost advantage but also increases (weakly, and perhaps strictly) the value of communication; thus the second source always (weakly) benefits the manufacturer.

and

$$S_2(y) = \begin{cases} \{(a,b) \in S | (a_m - a) \tan y < b - b_m\} & \text{if } 0 \le y \le y^\dagger, \\ \{(a,b) \in S | a_m - a < (b - b_m) \cot y\} & \text{if } y^\dagger < y < \pi - y^\dagger, \\ \{(a,b) \in S | (a_m - a) \tan y > b - b_m\} & \text{if } \pi - y^\dagger \le y \le \pi \end{cases} \tag{8}$$

where $y^\dagger = \tan^{-1} \frac{\bar{b}-\underline{b}}{\bar{a}-\underline{a}}$. The optimal total capacity conditional on each $S_i(y)$ is determined by equation $\Psi_i(k; y) = \min_{1 \le i \le n} \frac{r_i}{w - c_i}$ where $\Psi_i(k; y) = \mathbb{E}[\bar{F}((k-a)b) | S_i(y)]$. The $\Psi_i(k; y)$'s can be derived in closed form with tan and cot functions of y as presented in Appendix B.

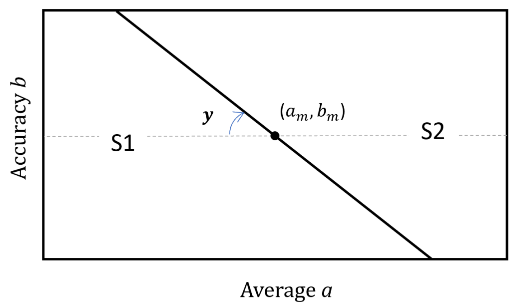

Figure 6: Type partition when (a, b) follows a continuous probability distribution $G(a, b)$.

With variable y, we can generalize our result that an informative equilibrium always exists. The intuition behind the proof is as follows. As y increases from 0 to π, subspaces S_1 and S_2 are flipped and so the optimal total capacities conditional on S_1 and S_2 are swapped. Because the optimal total capacity is continuous as a function of y, by the Intermediate Value Theorem, there must be some $y \in [0, \pi]$ such that the total capacity level conditional on S_1 is equal to that conditional on S_2. Furthermore, for $y \le \frac{\pi}{2}$, $S_1(y)$ represents a "low-average, low-accuracy" subspace and $S_2(y)$ represents a "high-average, high-accuracy" subspace. For $y > \frac{\pi}{2}$, $S_1(y)$ and $S_2(y)$ become "low-average, high-accuracy" and "high-average, low-accuracy," respectively.

Proposition 7. *When a and b are both continuous random variables with joint density function $g(a, b)$, there exists $y \in [0, \pi]$ such that an informative equilibrium exists in which for $i = 1, 2$, the retailer with private forecast $(a, b) \in S_i(y)$ reports $m = S_i(y)$ where $S_1(y)$ and $S_2(y)$ are as defined in* (7) *and* (8).

Proposition 8 proves that, as in the base model, the manufacturer's total capacity level in the informative equilibrium is identical to that without any communication. Consequently, informative communication will not make the retailer worse off as compared to the case of no communication but benefit the manufacturer by enabling a better informed decision on capacity reservation.

Proposition 8. *Under the informative equilibrium characterized in Proposition 7, the manufacturer's total capacity level is message-independent, and is also optimal to the manufacturer's problem without communication. Compared to the case without communication, informative and influential communication, whenever it occurs, will make the manufacturer better off while maintaining the retailer's payoff unchanged.*

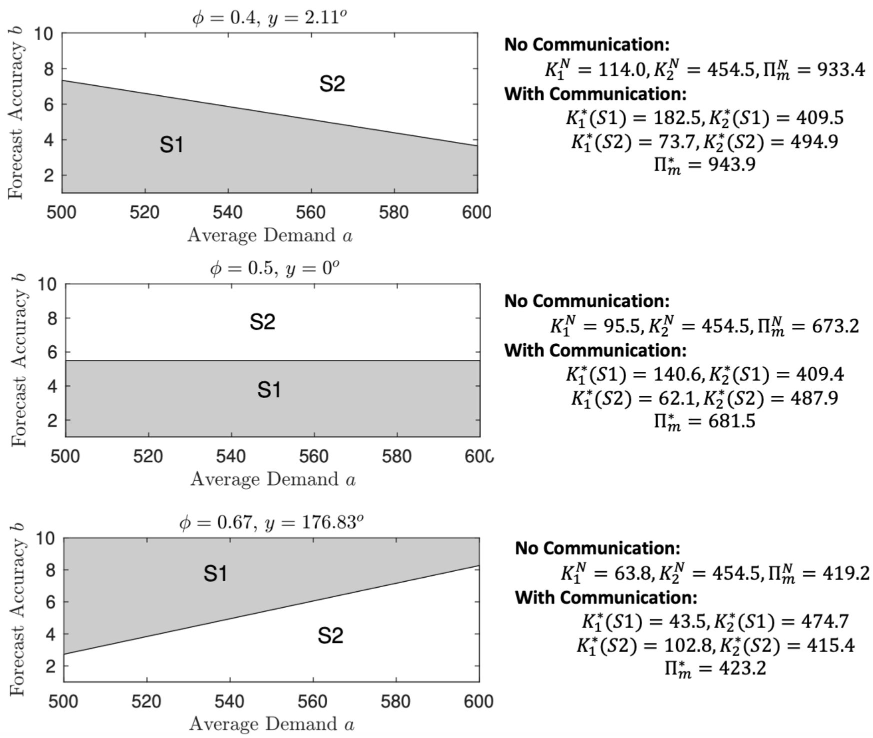

Figure 7: Illustration of informative partitions when (a, b) follows a joint continuous probability distribution.

We consider a set of numerical examples in which $a \sim U[500, 600]$, $b \sim U[1, 10]$ and $\epsilon \sim U[-500, 500]$ and they are independent. The manufacturer has two sources of capacity with $r_1 = 1, c_1 = 4.5$ and $r_2 = 5, c_2 = 0$. For each $w = 7, 6.5$ and 6 and correspondingly $\phi = 0.4, 0.5$ and 0.67, we numerically solve for the value of y leading to an informative equilibrium, and compute the equilibrium outcomes. The results are shown in Figure 7. Consistent with our findings from the base model, for the case of $\phi = 0.4$, the type space is partitioned into "low-average, low-accuracy" and "high-average, high-accuracy" regions. For the case of $\phi = 0.5$, the space is horizontally divided into two parts such that only accuracy information is revealed whereas the information about average demand remains as the prior conditional on each subspace. For $\phi = 0.67$, the type space is partitioned into "low-average, high-accuracy" and "high-average, low-accuracy" regions. Communication enables the manufacturer to reserve more flexible capacity and less inflexible capacity conditional on a high-accuracy message than conditional on a low-accuracy message, thereby improving the manufacturer's expected profit compared to the case without communication. Furthermore, in Appendix B, we numerically show that our findings regarding the impact of capacity cost reduction and expansion continue to hold when both a and b are continuous. The manufacturer can be worse off with a reduction in capacity costs or with portfolio expansion if these changes significantly attenuate information revelation compared to their operational benefit.

5.3 Informative Equilibria Supported by Asymmetric Partitions

In the base model, we have focused on symmetric partitions in the sense that the two partition points have equal distances from the middle point a_m, i.e., $x_l + x_h = 2a_m$. In general, informative communication can be supported by asymmetric partitions if $(x_l, x_h) \in [\underline{a}, \bar{a}]^2$ and the total capacity K_T constitute a solution to the system of nonlinear equations: $\phi = \Psi_1(K_T; x_l, x_h) = \Psi_2(K_T; x_l, x_h)$. Infinitely many (x_l, x_h) may exist to solve these equations since there are three unknowns with two equations.

For the special case of uniformly distributed a and ϵ, by fixing x_l or x_h at one of the endpoints of $[\underline{a}, \bar{a}]$, we are able to characterize a few representative asymmetric partitions under certain conditions.[17] Define $v_h = \frac{2\sigma\rho_l(b_h-b_l)}{b_h(\rho_l b_l+\rho_h b_h)}$ and $v_l = \frac{2\sigma\rho_h(b_h-b_l)}{b_l(\rho_l b_l+\rho_h b_h)}$.

Proposition 9. *Suppose that $a \sim U[\underline{a}, \bar{a}]$ and $\epsilon \sim U[-\sigma, \sigma]$ and that there are two capacity sources, neither of which dominates the other (i.e., $r_1 + c_1 > r_2 + c_2$ and $\frac{r_1}{w-c_1} < \frac{r_2}{w-c_2}$). The (x_l, x_h)'s given below support an informative and influential equilibrium if $\bar{a} - \underline{a} \geq |1 - 2\phi| \max(v_l, v_h)$, $\Psi_0(\underline{a} + \frac{\sigma}{b_h}) \leq \frac{r_1}{w-c_1}$, and $\min_{i=1,2} \Psi_i(\bar{a} - \frac{\sigma}{b_h}; x_l, x_h) \geq \frac{r_2-r_1}{c_1-c_2}$.*

(a) For $\phi \leq \frac{1}{2}$, $(x_l, x_h) = (\bar{a}, \underline{a} + (1-2\phi)v_h)$ and $(x_l, x_h) = (\bar{a} - (1-2\phi)v_l, \underline{a})$, and the manufacturer's equilibrium capacity levels satisfy $K_1^*(S_1) > K_1^*(S_2)$ and $K_2^*(S_1) < K_2^*(S_2)$.

(b) For $\phi > \frac{1}{2}$, $(x_l, x_h) = (\underline{a}, \bar{a} - (2\phi-1)v_h)$ and $(x_l, x_h) = (\underline{a} + (2\phi-1)v_l, \bar{a})$, and the manufacturer's equilibrium capacity levels satisfy $K_1^*(S_1) < K_1^*(S_2)$ and $K_2^*(S_1) > K_2^*(S_2)$.

The asymmetric partitions characterized in Proposition 9 are illustrated in Figure 8. These partitions and the corresponding equilibrium capacity levels exhibit similar properties as with the symmetric partitions derived earlier in Proposition 2. For the case of $\phi \leq 1/2$, we have $x_l > x_h$ and so

[17] The condition, $\bar{a} - \underline{a} \geq |1 - 2\phi| \max(v_l, v_h)$, ensures that x_l and x_h are within $[\underline{a}, \bar{a}]$. The other two conditions are analogous to the regularity conditions imposed in Proposition 2.

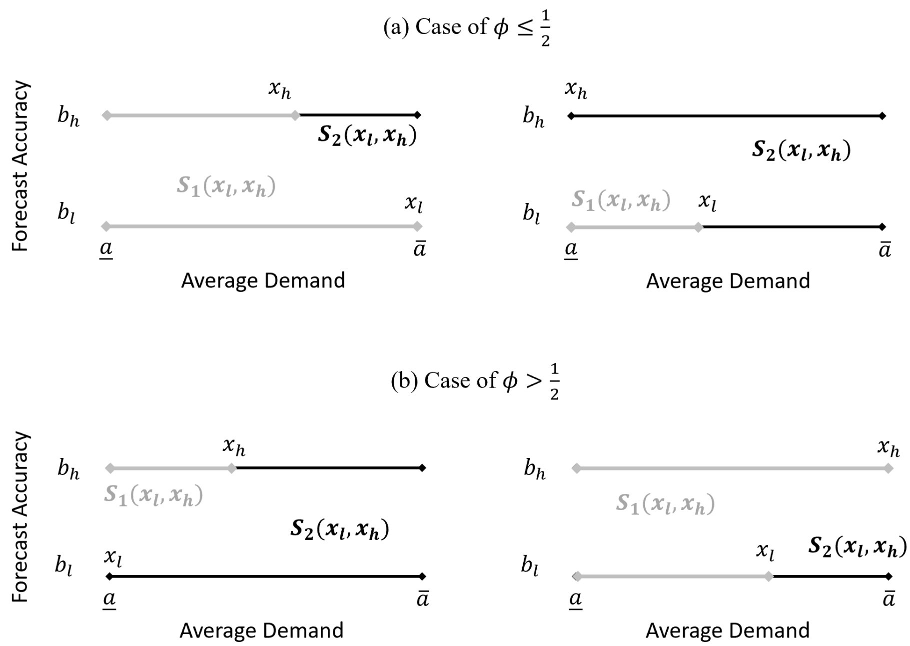

Figure 8: Illustration of asymmetric informative partitions

informative communication is supported by "low-average, low-accuracy" versus " high-average, high-accuracy" messages. For the case $\phi > 1/2$, $x_l < x_h$ and informative communication is supported by "low-average, high-accuracy" versus "high-average, low-accuracy" messages. The communication is also influential because the manufacturer reserves more flexible capacity source given a low-accuracy message than given a high-accuracy message.

Furthermore, it should be noted that the manufacturer's expected profit can vary under different forms of partitions. Suppose that the manufacturer can pinpoint a specific partition (and the retailer has no incentive not to conform during communication). Then, one can determine the profit-maximizing partition by optimizing the manufacturer's equilibrium profit, which is a function of $(x_l, x_h) \in [\underline{a}, \bar{a}]^2$, subject to a nonlinear equality constraint $\phi = \Psi_1(K_T^T; x_l, x_h) = \Psi_2(K_T^T; x_l, x_h)$. Since the retailer's profit is not impacted by communication, the equilibrium under the profit-maximizing partition is the most favorable one in the sense of Pareto-dominance. We conducted a numerical study on the profit-maximizing partition (x_l^*, x_h^*) and the resulting equilibrium. As reported in Appendix F, we observed that the profit-maximizing partition can be symmetric, asymmetric as characterized in Proposition 9 in which one of the x_l^* and x_h^* equals an endpoint of $[\underline{a}, \bar{a}]$, or of other asymmetric forms in between (i.e., neither symmetric nor with x_l^* or x_h^* at an endpoint).[18]

[18]Our numerical study suggests that for the special case of $\phi = 1/2$, the symmetric partition is the unique partition that supports an informative equilibrium and thus maximizes the manufacturer's profit. This is because given

Additionally, according to our numerical results, the maximized manufacturer profit is on average 1.37% higher than that under the symmetric partition. Hence, symmetric partitions can serve as a good approximation if it is difficult for supply chain firms to pinpoint and conform to a specific partition during communication. Nonetheless, all the equilibrium properties we derived earlier continue to hold under the profit-maximizing partition. For example, (i) consistent with Proposition 2, $x_l^* < x_h^*$ if $\phi > \frac{1}{2}$ whereas $x_l^* > x_h^*$ otherwise; (ii) when neither capacity sources dominates, the manufacturer tailors the capacity portfolio based on the revealed accuracy information; (iii) the difference $|x_h^* - x_l^*|$ and the value of communication attain their maximums at $\phi = \frac{1}{2}$ and shrink as ϕ moves further away from $\frac{1}{2}$ in either direction; and (iv) the manufacturer can be better off with an increase in the cost of the critical capacity source.

6 Conclusion

Motivated by the fact that some firms are now sharing information on forecast uncertainty and mean, we re-visit a canonical setting in supply chain research: A downstream party (the retailer) sources product, under a wholesale price contract, from an upstream party (the manufacturer) that must invest in capacity in advance of the retailer's order and then satisfies the order (subject to the capacity built) after demand is realized. Departing from the previous research, we consider both the forecast average and accuracy to be the retailer's private information. We establish that credible and informative communication emerges in equilibrium under very general conditions. Moreover, when the manufacturer has multiple sources of capacity that differ in reservation and execution costs, the communication can be influential, strictly improve the manufacturer's expected profit, and result in a Pareto-improvement of supply chain profits. Operational improvements to the manufacturer's capacity portfolio (e.g., cost reduction or expansion) can hurt the manufacturer if they attenuate information revelation during communication.

In the paper, we have focused on a specific supply portfolio in which capacity sources differ in reservation and execution costs. An interesting question for future research is whether cheap-talk communication of forecast mean and accuracy can be influential in other multi-sourcing settings (e.g., supply sources that differ in costs and responsiveness). More broadly, we hope our work will open up avenues for future research into multi-dimensional cheap talk in operations and supply

$\phi = 1/2$ (and a symmetric demand distribution), the manufacturer's total capacity is always equal to the mean demand conditional on any message, whereas any form of asymmetric partitions will entail different conditional means given different messages, thus making informative communication impossible.

chain settings with other types of private information beyond demand forecasts.

References

Allon, Gad, Achal Bassamboo. 2011. Buying from the babbling retailer? the impact of availability information on customer behavior. *Management Science* **57**(4) 713–726.

Allon, Gad, Achal Bassamboo, Itai Gurvich. 2011. "we will be right with you": Managing customer expectations with vague promises and cheap talk. *Operations research* **59**(6) 1382–1394.

Amazon Vendor Training Center. n.d. Vendsor central demand forecast: How to use the amazon retail analytics (ara) forecasting report. `https://s3.amazonaws.com/vendorcentral/VLC/North+Amercia/US/External/VOLT/33_Guide+your+demand+planning+with+Amazon's+forecasting+system/Optimize+inventory+levels+with+Vendor+Central+Demand+Forecast.pdf` accessed on Feb 28, 2024.

Ang, Erjie, Dan A Iancu, Robert Swinney. 2017. Disruption risk and optimal sourcing in multitier supply networks. *Management Science* **63**(8) 2397–2419.

Avinadav, Tal, Noam Shamir. 2023. Partial vertical ownership, capacity investment, and information exchange in a supply chain. *Management Science* .

Battaglini, Marco. 2002. Multiple referrals and multidimensional cheap talk. *Econometrica* **70**(4) 1379–1401.

Beer, Ruth, Anyan Qi. 2022. To communicate or not? inter-firm communication in collaborative projects. *Managememt Science, Forthcoming* .

Berman, Oded, Mohammad M Fazel-Zarandi, Dmitry Krass. 2019. Truthful cheap talk: Why operational flexibility may lead to truthful communication. *Management Science* **65**(4) 1624–1641.

Cachon, Gérard P, Martin A Lariviere. 2001. Contracting to assure supply: How to share demand forecasts in a supply chain. *Management science* **47**(5) 629–646.

Celikbas, Muruvvet, George J Shanthikumar, Jayashankar M Swaminathan. 1999. Coordinating production quantities and demand forecasts through penalty schemes. *IIE Transactions* **31**(9) 851–864.

Chakraborty, Archishman, Rick Harbaugh. 2007. Comparative cheap talk. *Journal of Economic Theory* **132**(1) 70–94.

Chakraborty, Archishman, Rick Harbaugh. 2010. Persuasion by cheap talk. *American Economic Review* **100**(5) 2361–82.

Chopra, Sunil. 2019. *Supply chain management: Strategy, planning and operation*. 7th ed. Pearson, New York, NY.

Chu, Leon Yang, Noam Shamir, Hyoduk Shin. 2017. Strategic communication for capacity alignment with pricing in a supply chain. *Management Science* **63**(12) 4366–4388.

Cohen, Morris A, Teck H Ho, Z Justin Ren, Christian Terwiesch. 2003. Measuring imputed cost in the semiconductor equipment supply chain. *Management Science* **49**(12) 1653–1670.

Crawford, Vincent P, Joel Sobel. 1982. Strategic information transmission. *Econometrica: Journal of the Econometric Society* 1431–1451.

Dada, Maqbool, Nicholas C Petruzzi, Leroy B Schwarz. 2007. A newsvendor's procurement problem when suppliers are unreliable. *Manufacturing & Service Operations Management* **9**(1) 9–32.

Dong, Lingxiu, Xin Geng, Guang Xiao, Nan Yang. 2022. Procurement strategies with unreliable suppliers under correlated random yields. *Manufacturing & service operations management* **24**(1) 179–195.

Donohue, Karen, Elena Katok, Stephen Leider. 2018. The handbook of behavioral operations .

Engardio, Pete. 2001. Why the supply chain broke down. *Business Week* **3724** 41.

Feng, Qi, Guoming Lai, Lauren Xiaoyuan Lu. 2015. Dynamic bargaining in a supply chain with asymmetric demand information. *Management Science* **61**(2) 301–315.

Fu, Qi, Chung-Yee Lee, Chung-Piaw Teo. 2010. Procurement management using option contracts: random spot price and the portfolio effect. *IIE transactions* **42**(11) 793–811.

Hu, Bin, Izak Duenyas, Damian R Beil. 2013. Does pooling purchases lead to higher profits? *Management Science* **59**(7) 1576–1593.

Intentwise, Inc. n.d. Highlights from amazon demand forecast report. `https://www.intentwise.com/analytics-cloud/data-source/amazon-vendor/demand-forecast` accessed on Feb 21, 2024.

Kolassa, Stephan, Bahman Rostami-Tabar, Enno Siemsen. 2023. *Demand Forecasting for Executives and Professionals*. CRC Press.

Li, Xiaolin, Özalp Özer, Upender Subramanian. 2022. Are we strategically naïve or guided by trust and trustworthiness in cheap-talk communication? *Management Science* **68**(1) 376–398.

Lu, Tao. 2024. Can a supplier's yield risk be truthfully communicated via cheap talk? *Manufacturing & Service Operations Management* Forthcoming.

Lu, Tao, Brian Tomlin. 2022. Sourcing from a self-reporting supplier: Strategic communication of social responsibility in a supply chain. *Manufacturing & Service Operations Management* **24**(2) 902–920.

Martínez-de Albéniz, Victor, David Simchi-Levi. 2005. A portfolio approach to procurement contracts. *Production and Operations Management* **14**(1) 90–114.

Martinez-de Albeniz, Victor, David Simchi-Levi. 2009. Competition in the supply option market. *Operations Research* **57**(5) 1082–1097.

Niranjan, Tarikere T, Shobhit Mathur, Narendra Kumar Ghosalya, Nagesh Gavirneni. 2022. Inflating orders can worsen supply chain uncertainty. *Harvard Business Review (Digital Article, October 27)* .

Özer, Özalp, Wei Wei. 2006. Strategic commitments for an optimal capacity decision under asymmetric forecast information. *Management science* **52**(8) 1238–1257.

Özer, Özalp, Yanchong Zheng, Kay-Yut Chen. 2011. Trust in forecast information sharing. *Management Science* **57**(6) 1111–1137.

Ren, Z Justin, Morris A Cohen, Teck H Ho, Christian Terwiesch. 2010. Information sharing in a long-term supply chain relationship: The role of customer review strategy. *Operations research* **58**(1) 81–93.

Shamir, Noam, Hyoduk Shin. 2016. Public forecast information sharing in a market with competing supply chains. *Management Science* **62**(10) 2994–3022.

Sobel, Joel. 2020. Signaling games. *Complex Social and Behavioral Systems: Game Theory and Agent-Based Models* 251–268.

Taylor, Terry A, Wenqiang Xiao. 2010. Does a manufacturer benefit from selling to a better-forecasting retailer? *Management Science* **56**(9) 1584–1598.

Terwiesch, Christian, Z Justin Ren, Teck H Ho, Morris A Cohen. 2005. An empirical analysis of forecast sharing in the semiconductor equipment supply chain. *Management science* **51**(2) 208–220.

Online Appendices for "Credible Cheap-Talk Communication of Private Demand Information on Both the Forecast Average and Accuracy"

A Supplemental Materials for the Case of More Than Two Capacity Sources

In this appendix, we present some additional results for the case of n capacity sources which are omitted from §5.1 of the main paper.

A.1 Problem Formulation

Receiving any message m, the manufacturer determines a vector $\mathbf{K} = (K_1, K_2, ..., K_n)$ to maximize the following expected profit, where K_i is the amount of capacity source i to be reserved.

$$\Pi_m(\mathbf{K}|m) = \mathbb{E}[R(\mathbf{K}, D)|m] - \sum_{i=1}^{n} r_i K_i. \tag{A.1}$$

In the above, $R(\mathbf{K}, D)$ is the optimal objective value of the second-stage problem where the manufacturer executes the reserved capacity after demand realization:

$$\begin{aligned} R(\mathbf{K}, D) = \quad \max_{x_1, x_2, \ldots, x_n} \quad & w \min\left(\textstyle\sum_{i=1}^{n} x_i, D\right) - \textstyle\sum_{i=1}^{n} c_i x_i \\ s.t. \quad & 0 \leq x_i \leq K_i \text{ for } i = 1, 2, ..., n. \end{aligned} \tag{A.2}$$

Similar to our base model, the retailer's expected profit depends on the manufacturer's capacity reservation only through the total capacity level $\sum_{i=1}^{n} K_i$. The retailer's expected profit is given by

$$\Pi_r(\mathbf{K}) = (p - w)\mathbb{E}[\min(\sum_{i=1}^{n} K_i, D)]. \tag{A.3}$$

A.2 The Manufacturer's Optimal Portfolio

The following lemma provides a procedure to determine the manufacturer's optimal capacity levels $\mathbf{K}^*$ based on any posterior demand distribution $F_D(\cdot|m)$. In the optimal solution, the manufacturer may procure (reserve) a subset of the available capacity sources. Algorithm A.1 can be used to determine the optimal portfolio of procured sources, denoted by I^*, and then the optimal capacity levels can be computed based on the first-order conditions.

Lemma A.1. *Given any message m, the set of capacity sources used by the manufacturer in the optimal solution I^* and the optimal capacity levels $\mathbf{K}^*$ can be determined by the following.*

Algorithm A.1. Find $i_A = \arg\min_{1\leq i\leq n} \frac{r_i}{w-c_i}$ and $i_B = \arg\min_{1\leq i\leq n}(r_i + c_i)$.[19] Set $I^* = \{i_B\}$.

Step 1. If $i_A = i_B$, stop and output I^*. Otherwise, set $i' = i_B$ and go to Step 2.

Step s = 2, ... Find $j = \arg\max_{i_A\leq i<i'} \frac{r_{i'}-r_i}{c_i-c_{i'}}$. Set $I^* = I^* \bigcup\{j\}$. If $j = i_A$, stop and output I^*; otherwise, set $i' = j$ and go to Step $s + 1$.

For all $i \notin I^*$, $K_i^* = 0$. For $i \in I^*$, (i) if $|I^*| = 1$, then $K^*_{i_A} = \bar{F}_D^{-1}(\frac{r_{i_A}}{w-c_{i_A}}|m)$; (ii) if $|I^*| = n^* \geq 2$, then let (j) denote the j-th smallest index in I and the optimal capacity levels are determined by

$$K_{(n^*)} = \bar{F}_D^{-1}\left(\frac{r_{(n^*)} - r_{(n^*-1)}}{c_{(n^*-1)} - c_{(n^*)}}\middle| m\right)$$

and

$$K_{(j)} = \bar{F}_D^{-1}\left(\frac{r_{(j)} - r_{(j-1)}}{c_{(j-1)} - c_{(j)}}\middle| m\right) - \bar{F}_D^{-1}\left(\frac{r_{(j+1)} - r_{(j)}}{c_{(j)} - c_{(j+1)}}\middle| m\right) \text{ for } j = n^* - 1, ..., 1$$

where $r_{(0)} = 0$ and $c_{(0)} = w$.

In particular, the manufacturer's optimal total capacity $K_T^* = \sum_{i=1}^n K_i^* = \bar{F}_D^{-1}(\frac{r_{i_A}}{w-c_{i_A}}|m)$.

Note that in the main paper, source i_A in the above algorithm is denoted by κ and referred to as the critical source. To provide a high-level idea of Algorithm A.1, we first find the source with the lowest ratio $\frac{r_i}{w-c_i}$ (thus the one providing the highest service level) and also the source with the lowest overall unit cost $r_i + c_i$, i.e., sources i_A and i_B respectively. If the two sources happen to coincide, then it is optimal to use this single exclusively. Otherwise, starting from i_B, we search among more flexible sources and include a source if it helps improve the service level until reaching i_A. Hence, the optimal portfolio I^* leverages the sources with the lowest possible overall costs while maintaining as a high service level as possible. We note that Fu et al. (2010) propose a similar solution approach for a manufacturer selecting the optimal portfolio of supply option contracts without explicitly identifying the two important capacity sources, i_A and i_B. In our problem, the optimal capacity portfolio will serve as the manufacturer's best response to any retailer message m, and as discussed in the main paper, source i_A plays a crucial role in determining the information revelation in equilibrium.

[19] We can break a tie by selecting the smaller index.

A.3 Additional Numerical Results for the Case of More Than Two Capacity Sources

In Proposition 4 of §4.3, we have shown that in the case of two capacity sources available, the manufacturer's expected profit can be increasing in the reservation and execution costs of the flexible source 1. In other words, the manufacturer may be worse off with a reduction in capacity costs. The reason is that a reduction in capacity cost can reduce information revelation during communication, and this negative impact may dominate the operational benefit of cost reduction. In fact, Proposition 4 is proved under a general case with n sources being reserved in the optimal portfolio. The general result is stated below.

Proposition A.1. *Consider the case of $a \sim U[\underline{a}, \bar{a}]$ and $\epsilon \sim [-\sigma, \sigma]$. Suppose that $\Psi_0(\underline{a}+\frac{\sigma}{b_h}) \leq \frac{r_1}{w-c_1}$ and $\min_{i=1,2} \Psi_i(\bar{a} - \frac{\sigma}{b_h}; x_l, x_h) \geq \frac{r_n - r_{n-1}}{c_{n-1}-c_n}$. If $n \geq 2$ capacity sources, indexed by $1, 2, ..., n$, are reserved in the optimal portfolio, then, for any change in capacity costs that does not affect the optimal portfolio selection, the following statements hold. (i) The manufacturer's expected profit is decreasing in r_i and c_i for all $i \geq 2$. (ii) The manufacturer's expected profit can be increasing in r_1 within an interval $\phi \in [\frac{1}{2} - \delta, \frac{1}{2}]$ if*

$$\bar{a} - \underline{a} < \frac{(b_h - b_l)^3 \sigma \sum_{i=2}^{n} (2z_i - 1)^2}{b_h b_l \left((b_h - b_l)^2 \sum_{i=2}^{n} (2z_i - 1) + (2z_2 - 1)(b_h + b_l)^2 (w - c_1)\right)}, \tag{A.4}$$

and can be increasing in c_1 within an interval $\phi \in [\frac{1}{2} - \delta, \frac{1}{2}]$ if

$$\bar{a} - \underline{a} < \frac{2(b_h - b_l)^3 \sigma \sum_{i=2}^{n} (2z_i - 1)^2}{b_h b_l (2(b_h - b_l)^2 \sum_{i=2}^{n} (2z_i - 1) + (b_h + b_l)^2 (4z_2^2 - 1)(w - c_1))} \tag{A.5}$$

where $z_i = \frac{r_i - r_{i-1}}{c_{i-1} - c_i}$ for $i = 2, ..., n$.

Proof. The proof of this proposition is contained in that of Proposition 4. □

Consider a numerical example with the same parameter values as in Example 1 in §5.1 of the main paper. With all other parameters fixed, we compute the manufacturer's expected profits while varying source 1's reservation cost r_1 and execution cost c_1. The left panel of Figure A.1 plots the manufacturer's expected profit with communication Π_m^* and its expected profit without communication Π_m^N, as r_1 varies from 0.3 to 0.9 with c_1 fixed as 5.0, whereas the right panel plots these profits as execution cost c_1 varies from 4.5 to 5.2 with r_1 fixed as 0.7. The profit is always higher with communication. Without communication, as expected, the manufacturer's profit Π_m^N is

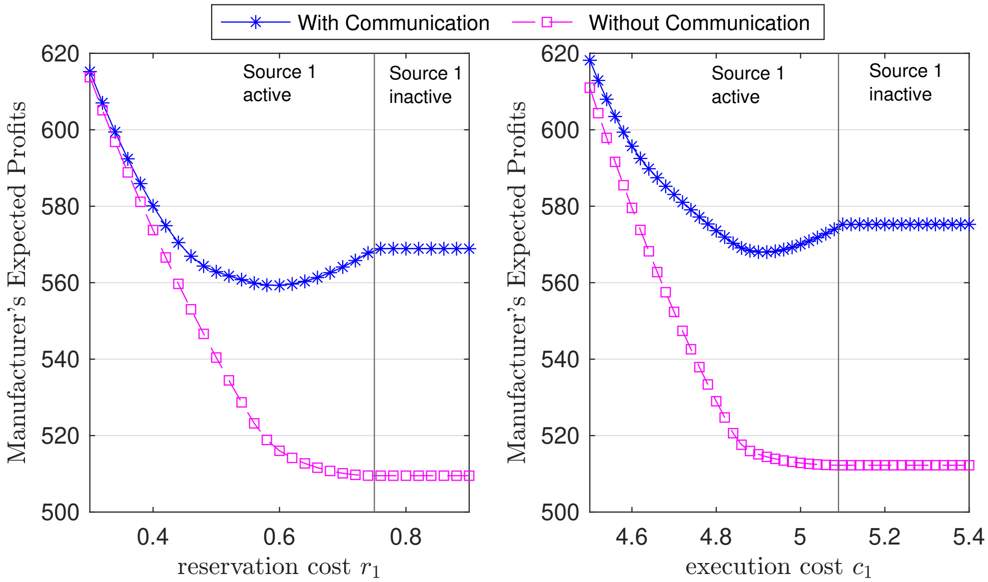

Figure A.1: The manufacturer's expected profits with and without communication

decreasing in r_1. However, with communication, Π_m^* is U-shaped in both r_1 and c_1 when source 1 is active (i.e., reserved). That is, in this example, a marginal reduction in reservation or execution cost decreases the manufacturer's profit unless these costs are already sufficiently low. This U-shaped profit is a result of the competing tradeoff between the operational advantage of cost reduction and the disadvantage of diminished information revelation.

B Supplemental Results for the Case of Continuous (a, b)

B.1 Closed-Form Expressions of the $\Psi_i(k; y)$

Let $y^\dagger = tan^{-1}\frac{\bar{b}-\underline{b}}{\bar{a}-\underline{a}}$. With tan and *cot* functions, the $\Psi_i(k;y)$'s can be written in closed form for $y \in [0, \pi]$ as follows.

$$\Psi_1(k;y) = \begin{cases} \frac{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{(a_m-a)\tan y + b_m} \bar{F}((k-a)b)g(a,b)dbda}{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{(a_m-a)\tan y + b_m} g(a,b)dbda} & \text{if } y \in [0, y^\dagger] \\ \frac{\int_{\underline{b}}^{\bar{b}} \int_{\underline{a}}^{a_m-(\bar{b}-b_m)\cot y} \bar{F}((k-a)b)g(a,b)dadb}{\int_{\underline{b}}^{\bar{b}} \int_{\underline{a}}^{a_m-(b-b_m)\cot y} g(a,b)dadb} & \text{if } y \in (y^\dagger, \pi - y^\dagger) \\ \frac{\int_{\underline{a}}^{\bar{a}} \int_{(a_m-a)\tan y + b_m}^{\bar{b}} \bar{F}((k-a)b)g(a,b)dbda}{\int_{\underline{a}}^{\bar{a}} \int_{(a_m-a)\tan y + b_m}^{\bar{b}} g(a,b)dbda} & \text{if } y \in [\pi - y^\dagger, \pi] \end{cases} \quad \text{(B.1)}$$

and

$$\Psi_2(k;y) = \begin{cases} \dfrac{\int_{\underline{a}}^{\bar{a}} \int_{(a_m-a)\tan y+b_m}^{\bar{b}} \bar{F}((k-a)b)g(a,b)dbda}{\int_{\underline{a}}^{\bar{a}} \int_{(a_m-a)\tan y+b_m}^{\bar{b}} g(a,b)dbda} & \text{if } y \in [0, y^\dagger] \\ \dfrac{\int_{\underline{b}}^{\bar{b}} \int_{a_m-(b-b_m)\cot y}^{\bar{a}} \bar{F}((k-a)b)g(a,b)dadb}{\int_{\underline{b}}^{\bar{b}} \int_{a_m-(b-b_m)\cot y}^{\bar{a}} g(a,b)dadb} & \text{if } y \in (y^\dagger, \pi - y^\dagger) \\ \dfrac{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{(a_m-a)\tan y+b_m} \bar{F}((k-a)b)g(a,b)dadb}{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{(a_m-a)\tan y+b_m} g(a,b)dadb} & \text{if } y \in [\pi - y^\dagger, \pi] \end{cases} \tag{B.2}$$

The derivation of these expressions is based on an idea as illustrated in Figure B.1.

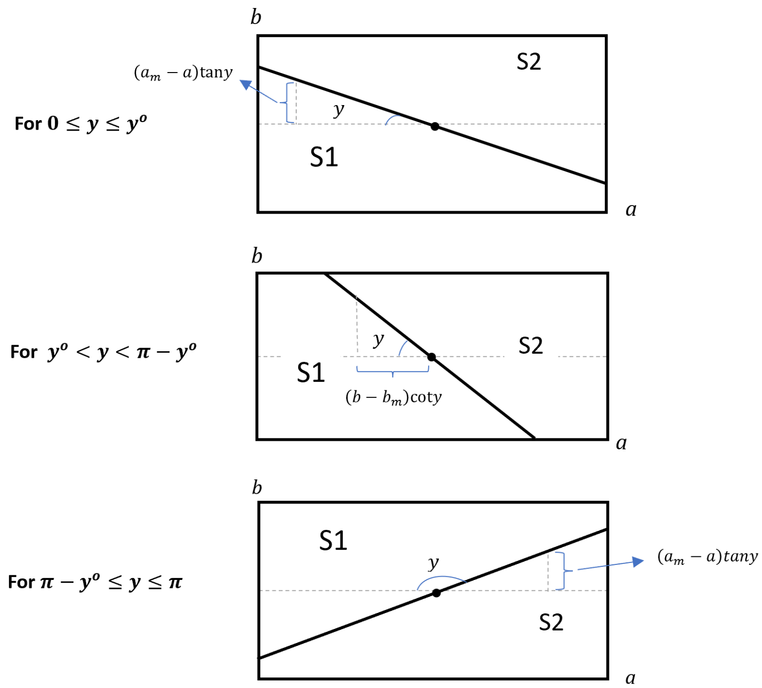

Figure B.1: Illustration of the subspace definition.

B.2 Impact of Capacity Cost Reduction and Expansion

In this appendix, we present a set of numerical examples to show that our findings as to the implications of portfolio cost reduction and expansion are robust when both demand average and forecast accuracy are continuous random variables.

We assume that a and b are independent and uniformly distributed: $a \sim U[500, 510]$ and $b \sim U[1, 10]$. The noise term ϵ follows a truncated normal distribution over $[-500, 500]$ with mean 0 and standard deviation 1000. The wholesale price is $w = 6.5$. Consider three capacity sources $n = 3$ with $(r_2, c_2) = (1, 4.5)$, $(r_3, c_3) = (5, 0)$, and (r_1, c_1) being varied as indicated in Figure B.2. In the example, source 2, if being used in isolation, provides an optimal service level of 0.5. source 1 will be used as a critical source only if $r_1 < 0.75$ in Panel (a) or $c_1 < 5.1$ in Panel (b). Similar to

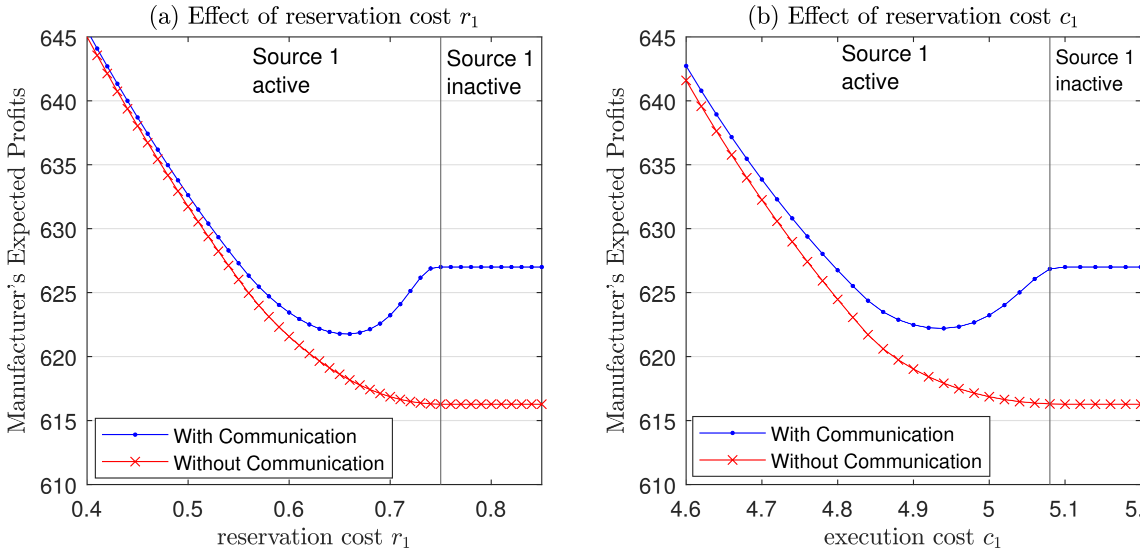

Figure B.2: The manufacturer's expected profits as functions of r_1 and c_1 when $a \sim U[500, 510]$ and $b \sim U[1, 10]$. In Panel (a), r_1 varies with $c_1 = 5$; In Panel (b), c_1 varies with $r_1 = 0.7$

the findings with the base model, the manufacturer's expected profit is U-shaped in r_1 and c_1, as the increases in r_1 and c_1 bring $\frac{r_1}{w-c_1}$ closer to 1/2, thus maximizing the value of communication. Comparing against the profits when source 1 is not active (i.e., not reserved), we can see that having source 1 active may lead to a lower profit. Thus, expanding the capacity portfolio with one additional source can possibly hurt the manufacturer if the additional source serves as the critical source resulting in an optimal service level further away from 0.5 (which attenuates the value of communication).

C Proofs

Proof of Lemma 1. First, consider the second-stage problem (2) for any realized D and given K_1 and K_2. Since $w > c_1 > c_2$, it is optimal to first execute the capacity from source 2 and then capacity from source 1, until demand is fulfilled or the total capacity is saturated. The optimal executed quantities are given by $x_2^* = \min(K_2, D)$ and $x_1^* = \min(K_1 + K_2, D) - x_2^*$. Substituting x_1^* and x_2^* into the first-stage problem (1) and rearranging the terms, we can simplify the objective function to

$$\Pi_m(K_1, K_2) = (w - c_2)\mathbb{E}[D|m] - (w - c_1)\mathbb{E}[(D - K_1 - K_2)^+|m] - (c_1 - c_2)\mathbb{E}[(D - K_2)^+|m] - \sum_{i=1}^{2} r_i K_i,$$

where $\mathbb{E}[\cdot|m]$ represents the expectation over random demand D conditional on message m. $\frac{\partial \Pi_m(K_1,K_2)}{\partial K_1} = (w-c_1)\Pr(D > K_1+K_2|m) - r_1$ and $\frac{\partial \Pi_m(K_1,K_2)}{\partial K_2} = (w-c_1)\Pr(D > K_1+K_2|m) + (c_1-c_2)\Pr(D > K_2|m) - r_2$. The characterization of K_1^* and K_2^* follows by a case-by-case analysis on the optimality conditions: For $i = 1, 2$, if $K_i^* > 0$, $\frac{\partial \Pi_m(K_1^*,K_2^*)}{\partial K_i} = 0$; if $K_i^* = 0$, $\frac{\partial \Pi_m(K_1^*,K_2^*)}{\partial K_i} \leq 0$. $\square$

Proof of Lemma 2. The first part, $K_T^*(S_1) = K_T^*(S_2)$, follows immediately by (5) as discussed in the main paper. For the second part, note that the total capacity K_T is determined by an equation $\Psi_i(K_T; x_l, x_h) =$ some constant, which is $\frac{r_1}{w-c_1}$ for cases (1) (3) of Lemma 1 or $\frac{r_2}{w-c_2}$ for case (2) of Lemma 1. We consider the constant to be $\frac{r_1}{w-c_1}$ in the following proof (which also applies to the case of $\frac{r_2}{w-c_2}$). It suffices to show that in general any K_T that solves (5) coincides with the manufacturer's optimal total capacity K_T^N based on the prior belief of (a, b). By Proposition 1, K_T^N is determined by the equation $\frac{r_1}{w-c_1} = \Psi_0(K_T^N)$ where

$$\Psi_0(k) = \rho_l \int_{\underline{a}}^{\bar{a}} \bar{F}((k-a)b_l)g(a)da + \rho_h \int_{\underline{a}}^{\bar{a}} \bar{F}((k-a)b_h)g(a)da$$

Consider any solution (x_l, x_h, K_T) to (5). We must have

$$(\rho_l G(x_l) + \rho_h G(x_h))\frac{r_1}{w-c_1} = \rho_l \int_{\underline{a}}^{x_l} \bar{F}((K_T-a)b_l)g(a)da + \rho_h \int_{\underline{a}}^{x_h} \bar{F}((K_T-a)b_h)g(a)da$$

and

$$(\rho_l \bar{G}(x_l) + \rho_h \bar{G}(x_h))\frac{r_1}{w-c_1} = \rho_l \int_{x_l}^{\bar{a}} \bar{F}((K_T-a)b_l)g(a)da + \rho_h \int_{x_h}^{\bar{a}} \bar{F}((K_T-a)b_h)g(a)da.$$

Since $\rho_l + \rho_h = 1$ and $G(x) + \bar{G}(x) = 1$, summing up both sides of the above two equations yields

$$\frac{r_1}{w-c_1} = \rho_l \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T-a)b_l)g(a)da + \rho_h \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T-a)b_h)g(a)da.$$

Therefore, K_T also solves $\frac{r_1}{w-c_1} = \Psi_0(K_T)$. Thus, K_T is an optimal solution to the manufacturer's problem without communication. $\square$

Proof of Proposition 1. By Lemma 1, regardless of whether there is a dominating capacity source, given $(a, b) \in S_i(x_l, x_h)$, the total capacity K_T is determined by an equation $\Psi_i(K_T; a_m - y, a_m + y) =$ some constant, which is $\frac{r_1}{w-c_1}$ for cases (1) (3) of Lemma 1 or $\frac{r_2}{w-c_2}$ for case (2) of Lemma 1. It suffices to show that for any given K_T, there is a $y \in [-\frac{\bar{a}-\underline{a}}{2}, \frac{\bar{a}-\underline{a}}{2}]$ such that $\Psi_1(K_T; a_m - y, a_m +$

$y) = \Psi_2(K_T; a_m - y, a_m + y)$. Define $\Delta(y) = \Psi_1(K_T; a_m - y, a_m + y) - \Psi_2(K_T; a_m - y, a_m + y)$. Note that at $y = -\frac{\bar{a}-\underline{a}}{2}$, $x_l = \bar{a}$ and $x_h = \underline{a}$ while at $y = \frac{\bar{a}-\underline{a}}{2}$, $x_l = \underline{a}$ and $x_h = \bar{a}$. Moreover, notice that $\Psi_1(K_T, \bar{a}, \underline{a}) = \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_l)g(a)da$, $\Psi_2(K_T, \bar{a}, \underline{a}) = \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_h)g(a)da$, $\Psi_1(K_T, \underline{a}, \bar{a}) = \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_h)g(a)da$, and $\Psi_2(K_T, \underline{a}, \bar{a}) = \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_l)g(a)da$. Hence,

$$\Delta(-\frac{\bar{a}-\underline{a}}{2}) = \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_l)g(a)da - \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_h)g(a)da$$

and

$$\Delta(\frac{\bar{a}-\underline{a}}{2}) = \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_h)g(a)da - \int_{\underline{a}}^{\bar{a}} \bar{F}((K_T - a)b_l)g(a)da$$

Therefore, $\Delta(y)$ must have opposite signs at $y = -\frac{\bar{a}-\underline{a}}{2}$ and $y = \frac{\bar{a}-\underline{a}}{2}$. Since $\Delta(y)$ is continuous, by the Intermediate Value Theorem, there exists a $y \in [-\frac{\bar{a}-\underline{a}}{2}, \frac{\bar{a}-\underline{a}}{2}]$ such that $\Delta(y) = 0$. Consequently, we can always find some $x_l = \frac{\underline{a}+\bar{a}}{2} - y$ and $x_h = \frac{\underline{a}+\bar{a}}{2} + y$ to satisfy (5) for any given K_T.

□

Proof of Proposition 2. We establish part (i) of the proposition in two steps.

Step 1. First, we show that y^o as given in (6) is such that $x_l = a_m - y^o$, $x_h = a_m + y^o$ and K_T satisfy (5), thus supporting an informative equilibrium. To get clean closed-form expressions of K_T and y^o, we assume $(K_T^N - a)b \in [-\sigma, \sigma]$ for all $(a, b) \in S$, i.e., $\bar{a} - \frac{\sigma}{b_h} \leq K_T^N \leq \underline{a} + \frac{\sigma}{b_h}$, so that we can substitute $\bar{F}(x) = \frac{\sigma - x}{2\sigma}$ into the integrals without considering the nasty cases where x is out of $[-\sigma, \sigma]$. Because $\Psi_0(k)$ is decreasing, K_T^N will be in the desired range if $\Psi_0(\underline{a} + \frac{\sigma}{b_h}) \leq \frac{r_1}{w - c_1}$ and $\Psi_0(\bar{a} - \frac{\sigma}{b_h}) \geq \frac{r_1}{w - c_1}$: The former inequality is exactly the first regularity condition, while the latter inequality is implied by the second regularity condition $\min_{i=1,2}\{\Psi_i(\bar{a} - \frac{\sigma}{b_h}; a_m - y^o, a_m + y^o)\} \geq \frac{r_2 - r_1}{c_1 - c_2}$. This is because (1) $\frac{r_2 - r_1}{c_2 - c_1} > \frac{r_1}{w - c_1}$ in the case of no dominating capacity source (i.e., $r_2 + c_2 < r_1 + c_1$ and $\frac{r_1}{w - c_1} < \frac{r_2}{w - c_2}$), and (2) $\Psi_0(\bar{a} - \frac{\sigma}{b_h}) = (\rho_l G(x_l) + \rho_h G(x_h))\Psi_1(\bar{a} - \frac{\sigma}{b_h}; a_m - y^o, a_m + y^o) + (\rho_l \bar{G}(x_l) + \rho_h \bar{G}(x_h))\Psi_2(\bar{a} - \frac{\sigma}{b_h}; a_m - y^o, a_m + y^o) \geq \min_i \Psi_i(\bar{a} - \frac{\sigma}{b_h}; a_m - y^o, a_m + y^o)$.

Let $\phi = \frac{r_1}{w - c_1}$. Under these conditions, we can derive $K_T^N = \frac{\underline{a}+\bar{a}}{2} + \frac{\sigma(1-2\phi)}{\rho_h b_h + \rho_l b_l}$. From Lemma 2, we also have $K_T = K_T^N$. Then, solving the equation $\Psi_1(K_T^N; a_m - y, a_m + y) = \Psi_2(K_T^N; a_m - y, a_m + y)$ where $a_m = \frac{\underline{a}+\bar{a}}{2}$, we find two roots:

$$y_1 = \frac{1}{2(\rho_l b_l + \rho_h b_h)^2} \left(4\sigma\rho_h\rho_l(b_h - b_l)(1 - 2\phi) - \sqrt{(\bar{a} - \underline{a})^2(\rho_l b_l + \rho_h b_h)^4 + 16\sigma^2\rho_h^2\rho_l^2(b_h - b_l)^2(1 - 2\phi)^2} \right)$$

and

$$y_2 = \frac{1}{2(\rho_l b_l + \rho_h b_h)^2}\left(4\sigma\rho_h\rho_l(b_h - b_l)(1-2\phi) + \sqrt{(\bar{a}-\underline{a})^2(\rho_l b_l + \rho_h b_h)^4 + 16\sigma^2\rho_h^2\rho_l^2(b_h-b_l)^2(1-2\phi)^2}\right).$$

From Proposition 1, we have known that there must exist a root in $[-\frac{\bar{a}-\underline{a}}{2}, \frac{\bar{a}-\underline{a}}{2}]$ such that $x_l = a_m - y^o$ and $x_h = a_m + y^o$ constitute a informative partition. We can eliminate one of the above roots based on whether $1 \geq 2\phi$ or $1 < 2\phi$.

(i) If $1 \geq 2\phi$, then it is easy to see that $y_2 > \frac{\bar{a}-\underline{a}}{2}$. So, let $y^o = y_1$. Moreover, in this case $y^o = y_1 < 0$ and hence $x_l > x_h$.

(ii) If $1 < 2\phi$, then it is easy to see $y_1 < -\frac{\bar{a}-\underline{a}}{2}$, so $y^o = y_2$. Moreover, in this case $y^o > 0$ and hence $x_l < x_h$.

Step 2. By Lemma 1, for both $i = 1, 2$, $K_2^*(S_i)$ is determined by equation $\frac{r_2-r_1}{c_1-c_2} = \Psi_i(K_2; x_l, x_h)$ and $\Psi_i(K_2; x_l, x_h)$ is decreasing in K_2. We focus on the cases where $(K_2^*(S_i) - a)b \in [-\sigma, \sigma]$ for all $(a, b) \in S$, or equivalently, $\bar{a} - \frac{\sigma}{b_h} \leq K_2^*(S_i) \leq \underline{a} + \frac{\sigma}{b_h}$. Since $K_2 \leq K_T$, all we need to assume is that $\bar{a} - \frac{\sigma}{b_h} \leq K_2$ and $K_T \leq \underline{a} + \frac{\sigma}{b_h}$. By the decreasing property of Ψ_0 and Ψ_i, this leads to the two regularity conditions $\Psi_i(\bar{a} - \frac{\sigma}{b_h}; x_l, x_h) \geq \frac{r_2-r_1}{c_1-c_2}$ for both $i = 1, 2$ and $\Psi_0(\underline{a} + \frac{\sigma}{b_h}) \leq \frac{r_1}{w-c_2}$.

Now we prove the relation of $K_2^*(S_1)$ and $K_2^*(S_2)$. Since $K_1^*(S_1) + K_2^*(S_1) = K_1^*(S_2) + K_2^*(S_2) = K_T^N$, the relation of $K_1^*(S_1)$ and $K_1^*(S_2)$ follows automatically. To this end, we examine the derivative of $\Delta(k) = \Psi_1(k; x_l, x_h) - \Psi_2(k; x_l, x_h)$ with respect to k.

$$\begin{aligned}\Delta'(k) = &-\frac{1}{\rho_l(x_l - \underline{a}) + \rho_h(x_h - \underline{a})}\left(b_l\rho_l\int_{\underline{a}}^{x_l} f((k-a)b_l)d_a + b_h\rho_h\int_{\underline{a}}^{x_h} f((k-a)b_l)d_a\right)\\ &+\frac{1}{\rho_l(\bar{a} - x_l) + \rho_h(\bar{a} - x_h)}\left(b_l\rho_l\int_{x_l}^{\bar{a}} f((k-a)b_l)d_a + b_h\rho_h\int_{x_h}^{\bar{a}} f((k-a)b_l)d_a\right)\end{aligned}$$

Under the regularity conditions, for all $k \in [\min(K_2^*(S_1), K_2^*(S_2)), K_T^N]$, $f((k-a)b) = \frac{1}{2\sigma}$ for all $(a, b) \in S$. Moreover, define $A = \frac{\bar{a}-\underline{a}}{2}$ and notice that $A - y^o = x_l - \underline{a} = \bar{a} - x_h$ and $A + y^o = \bar{a} - x_l = x_h - \underline{a}$. Substituting these relations and $f((k-a)b) = \frac{1}{2\sigma}$, we can simplify the above to

$$\Delta'(k) = \frac{-2\rho_h\rho_l(b_h - b_l)Ay^o}{\sigma(\rho_l(A - y^o) + \rho_h(A + y^o))(\rho_l(A + y^o) + \rho_h(A - y^o))}.$$

For the case of $\frac{r_1}{w-c_1} \leq \frac{1}{2}$, we have $y^o < 0$ and so $\Delta'(k) > 0$. Because $\Delta(K_T) = 0$, $\Delta(k) < 0$, that is, $\Psi_1(k; x_l, x_h) < \Psi_2(k; x_l, x_h)$ for $k \in [\min(K_2^*(S_1), K_2^*(S_2)), K_T^N]$. In the optimal solution, $\frac{r_2-r_1}{c_1-c_2} = \Psi_1(K_2^*(S_1); x_l, x_h) < \Psi_2(K_2^*(S_1); x_l, x_h)$. Since Ψ_2 is decreasing in k and $\frac{r_2-r_1}{c_1-c_2} =$

$\Psi_2(K_2^*(S_2); x_l, x_h)$, it follows that $K_2^*(S_1) < K_2^*(S_2)$.

Similarly, for the case of $\frac{r_1}{w-c_1} > \frac{1}{2}$, $y^o > 0$ and so $\Delta'(k) < 0$ and $\Psi_1(k; x_l, x_h) > \Psi_2(k; x_l, x_h)$ for $k \in [\min(K_2^*(S_1), K_2^*(S_2)), K_T^N]$. In the optimal solution, $\frac{r_2-r_1}{c_1-c_2} = \Psi_1(K_2^*(S_1); x_l, x_h) > \Psi_2(K_2^*(S_1); x_l, x_h)$. To have $\frac{r_2-r_1}{c_1-c_2} = \Psi_2(K_2^*(S_2); x_l, x_h)$ hold, we must have $K_2^*(S_1) > K_2^*(S_2)$.

For part (ii) of the proposition follows immediately by Proposition 1 and the fact that in any informative equilibrium the manufacturer chooses the same total capacity level, that is, either $K_1^*(m)$ or $K_2^*(m)$ in the case of only one capacity source being used, regardless of the retailer's message.

□

Proof of Corollary 1. We evaluate the derivative of y^o with respect to $(1-2\phi)$. The properties of $|x_h - x_l|$ follow since $|x_h - x_l| = 2|y^o|$. Define shorthand notation $B = \rho_l b_l + \rho_h b_h$ and $V = 4\sigma\rho_h\rho_l(b_h - b_l)$. For the case of $\phi \leq \frac{1}{2}$,

$$\frac{\partial y^o}{\partial(1-2\phi)} = \frac{V - V^2(1-2\phi)/\sqrt{(\bar{a}-\underline{a})^2B^4 + V^2}}{2B^2} \geq \frac{\phi V}{B^2} > 0$$

where the first inequality follows since $\bar{a} - \underline{a} > 0$ and $1 - 2\phi \geq 0$. Hence, y^o is increasing in $1-2\phi$, hence decreasing in ϕ for $\phi \in (0, \frac{1}{2}]$.

Similarly, for the case of $\phi > \frac{1}{2}$, we have

$$\frac{\partial y^o}{\partial(1-2\phi)} = \frac{V - V^2(2\phi-1)/\sqrt{(\bar{a}-\underline{a})^2B^4 + V^2}}{2B^2} > \frac{\phi V}{B^2} > 0$$

where the first inequality follows since $\bar{a} - \underline{a} > 0$ and $2\phi - 1 > 0$. Thus, y^o is decreasing in $\phi \in (\frac{1}{2}, 1)$.

By the expression of y^o, it is straightforward to verify that at $\phi = \frac{1}{2}$, $y^o = -\frac{\bar{a}-\underline{a}}{2}$ such that $|y^o|$ attains its maximum and $x_l = \bar{a}$ and $x_h = \underline{a}$. Furthermore, notice that

$$y^o \to \frac{V - \sqrt{(\bar{a}-\underline{a})^2B^4 + V^2}}{2B^2} \text{ as } \phi \to 0$$

and

$$y^o \to \frac{-V + \sqrt{(\bar{a}-\underline{a})^2B^4 + V^2}}{2B^2} \text{ as } \phi \to 1.$$

Therefore, by the monotone property of y^o characterized above and the fact that $y^o < 0$ (> 0) for $\phi \leq \frac{1}{2}$ $(> \frac{1}{2})$, $|y^o|$ attains the minimum value $\frac{\sqrt{(\bar{a}-\underline{a})^2B^4+V^2}-V}{2B^2}$, which is nonzero, as ϕ goes to 0 or 1.

□

Proof of Proposition 3. Let K_1^N and K_2^N represent the manufacturer's optimal capacity without any communication. Let S_1 and S_2 be any informative partitions. Clearly, $\Pi_m(K_1^N, K_2^N) = \sum_{i=1}^2 \Pr((a,b) \in S_i) \cdot \Pi_m(K_1^N, K_2^N | S_i) \leq \sum_{i=1}^2 \Pr((a,b) \in S_i) \cdot \Pi_m(K_1^*(S_i), K_2^*(S_i)|S_i)$ as $K_1^*(S_i)$ and $K_2^*(S_i)$ maximize Π_m conditional on $(a,b) \in S_i$ for each $i = 1, 2$. The inequality holds with equality if and only if K_i^N gives the same objective value as $K_i^*(S_1)$ and $K_i^*(S_2)$, i.e., multiple optimal solutions exist, for both $i = 1, 2$, However, when F is strictly increasing over its support, $K_2^*(S_i)$ is a unique solution to the optimality condition. Hence, the inequality holds strictly.

The retailer's payoff is determined solely by the manufacturer's total capacity level. In Lemma 2, we have shown that the total capacity $K_T^*(m)$ in any informative equilibrium is also optimal to the case without any communication. Thus, the retailer's payoff is unchanged. $\square$

Proof of Proposition 4. We prove this proposition for the case with n capacity sources being used in the optimal capacity portfolio. The characterization of the optimal capacity levels with n sources can be found in Lemma A.1 in Appendix A. We focus on relatively small changes in capacity costs without altering the optimal portfolio selection. Without loss of generality, we assume that all the n sources are active in the optimal portfolio (since one can simply remove inactive sources from the model and re-index the active ones). The statement of Proposition 4 follows by letting $n = 2$.

We first write out a general expression of the manufacturer's expected profit under the informative equilibrium.

$$\begin{aligned}
\Pi_m^* = (w - c_n)\mathbb{E}[D] &- \Pr(S_1)\left(\sum_{i=1}^n (c_{i-1} - c_i) \sum_{t=h,l} \frac{\rho_t G(x_t)}{\sum_{t'=h,l}\rho_t G(x_{t'})} \int_{\underline{a}}^{x_t}\int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=i}^n K_j^*(S_1))^+ \frac{f(\epsilon)g(a)}{G(x_t)} d\epsilon da\right) \\
&- \Pr(S_2)\left(\sum_{i=1}^n (c_{i-1} - c_i) \sum_{t=h,l} \frac{\rho_t \bar{G}(x_t)}{\sum_{t'=h,l}\rho_{t'} \bar{G}(x_{t'})} \int_{x_t}^{\bar{a}}\int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=i}^n K_j^*(S_2))^+ \frac{f(\epsilon)g(a)}{G(x_t)} d\epsilon da\right) \\
&- \Pr(S_1)\sum_{i=1}^n r_i K_i^*(S_1) - \Pr(S_2)\sum_{i=1}^n r_i K_i^*(S_2)
\end{aligned}$$

where $\Pr(S_1) = \sum_{t=h,l}\rho_t G(x_t)$ and $\Pr(S_2) = \sum_{t=h,l}\rho_t \bar{G}(x_t)$ denote the probabilities of $(a,b) \in S_1$ and $(a,b) \in S_2$, respectively.

In the case of $\rho_h = \rho_l = \frac{1}{2}$, the above profit can be simplified to

$$\Pi_m^* = \frac{(w - c_n)(\bar{a} + \underline{a})}{2} - \frac{\sum_{t=h,l}(L_{t1} + L_{t2})}{2} - \frac{\sum_{i=1}^n r_i \left(K_i^*(S_1) + K_i^*(S_2)\right)}{2} \tag{C.1}$$

where, for $t = h, l$,

$$L_{t1} = \sum_{i=1}^{n}(c_{i-1} - c_i) \int_{\underline{a}}^{x_t} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_1))^+ dF(\epsilon) dG(a)$$

and

$$L_{t2} = \sum_{i=1}^{n}(c_{i-1} - c_i) \int_{x_t}^{\bar{a}} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_2))^+ dF(\epsilon) dG(a).$$

Since the $\mathbf{K}^*(S_1)$ and $\mathbf{K}^*(S_2)$ are the optimal solutions maximizing the profits conditioning on S_1 and S_2, respectively, we can apply the Envelope Theorem when evaluating the derivatives of Π_m^*.

Impact of r_i and c_i $(i >= 2)$ For any noncritical source $i \geq 2$, the costs r_i and c_i will not impact x_l or x_h (provided that the change does not alter the portfolio selection). Hence, it is straightforward to show that, for $i \geq 2$,

$$\frac{\Pi_m^*}{\partial r_i} = -\frac{K_i^*(S_1) + K_i^*(S_2)}{2} \leq 0,$$

and

$$\frac{\Pi_m^*}{\partial c_i} = -\frac{1}{2} \sum_{t=h,l} \left(\frac{\partial L_{t1}}{\partial c_i} + \frac{\partial L_{t2}}{\partial c_i} \right) \leq 0$$

because, for all $t = l, h$,

$$\frac{\partial L_{t1}}{\partial c_i} = \int_{\underline{a}}^{x_t} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=i+1}^{n} K_j^*(S_1))^+ - (a + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_1))^+ dF(\epsilon) dG(a) \geq 0$$

and

$$\frac{\partial L_{t2}}{\partial c_i} = \int_{x_t}^{\bar{a}} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=i+1}^{n} K_j^*(S_2))^+ - (a + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_2))^+ dF(\epsilon) dG(a) \geq 0.$$

Impact of r_1 and c_1. As a change in r_1 or c_1 will impact $\phi = \frac{r_1}{w-c_1}$, thus affecting the values of x_l and x_h. The derivative of Π_m^* is much more involved.

$$\frac{\Pi_m^*}{\partial r_1} = -\frac{1}{2} \sum_{t=h,l} \left(\frac{\partial L_{t1}}{\partial r_1} + \frac{\partial L_{t2}}{\partial r_1} \right) - \frac{K_1^*(S_1) + K_1^*(S_2)}{2}$$

where

$$\frac{\partial L_{t1}}{\partial r_1} = \frac{\partial x_t}{\partial r_1} \sum_{i=1}^{n}(c_{i-1} - c_i) \int_{-\sigma}^{\sigma} (x_t + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_1))^+ g(x_t) dF(\epsilon)$$

and

$$\frac{\partial L_{t2}}{\partial r_1} = -\frac{\partial x_t}{\partial r_1} \sum_{i=1}^{n} (c_{i-1} - c_i) \int_{-\sigma}^{\sigma} (x_t + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_2))^+ g(x_t) dF(\epsilon).$$

Applying the fact that $\sum_{j=1}^n K_j^*(S_1) = \sum_{j=1}^n K_j^*(S_2)$, $\frac{\partial x_h}{\partial r_1} = \frac{\partial y^o}{\partial r_1}$ and $\frac{\partial x_l}{\partial r_1} = -\frac{\partial y^o}{\partial r_1}$, we have

$$\sum_{t=h,l} \left(\frac{\partial L_{t1}}{\partial r_1} + \frac{\partial L_{t2}}{\partial r_1} \right) = \frac{\partial y^o}{\partial r_1} \sum_{i=2}^{n} (c_{i-1} - c_i) \tau(\sum_{j=i}^{n} K_j^*(S_1), \sum_{j=i}^{n} K_j^*(S_2))$$

where

$$\begin{aligned} \tau(k_1, k_2) = & \int_{-\sigma}^{\sigma} (x_h + \frac{\epsilon}{b_h} - k_1)^+ g(x_h) - (x_l + \frac{\epsilon}{b_l} - k_1)^+ g(x_l) dF(\epsilon) \\ & - \int_{-\sigma}^{\sigma} (x_h + \frac{\epsilon}{b_h} - k_2)^+ g(x_h) - (x_l + \frac{\epsilon}{b_l} - k_2)^+ g(x_l) dF(\epsilon) \end{aligned}$$

Therefore, $\frac{\partial \Pi_m^*}{\partial r_1} > 0$ if and only if

$$\left(-\frac{\partial y^o}{\partial r_1} \right) \sum_{i=2}^{n} (c_{i-1} - c_i) tau(\sum_{j=i}^{n} K_j^*(S_1), \sum_{j=i}^{n} K_j^*(S_2)) > K_1^*(S_1) + K_1^*(S_2). \tag{C.2}$$

Now consider the derivative with respect to c_1. It follows that

$$\frac{\Pi_m^*}{\partial r_c} = -\frac{1}{2} \sum_{t=h,l} \left(\frac{\partial L_{t1}}{\partial c_1} + \frac{\partial L_{t2}}{\partial c_1} \right)$$

where

$$\begin{aligned} \frac{\partial L_{t1}}{\partial c_1} = & \int_{\underline{a}}^{x_t} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=2}^{n} K_j^*(S_1))^+ - (a + \frac{\epsilon}{b_t} - \sum_{j=1}^{n} K_j^*(S_1))^+ dF(\epsilon) dG(a) \\ & + \frac{\partial x_t}{\partial c_1} \sum_{i=1}^{n} (c_{i-1} - c_i) \int_{-\sigma}^{\sigma} (x_t + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_1))^+ g(x_t) dF(\epsilon) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial L_{t2}}{\partial c_1} = & \int_{x_t}^{\bar{a}} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=2}^{n} K_j^*(S_2))^+ - (a + \frac{\epsilon}{b_t} - \sum_{j=1}^{n} K_j^*(S_2))^+ dF(\epsilon) dG(a) \\ & - \frac{\partial x_t}{\partial c_1} \sum_{i=1}^{n} (c_{i-1} - c_i) \int_{-\sigma}^{\sigma} (x_t + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_2))^+ g(x_t) dF(\epsilon) \end{aligned}$$

Because $\frac{\partial x_h}{\partial c_1} = \frac{\partial y^o}{\partial c_1}$ and $\frac{\partial x_l}{\partial c_1} = -\frac{\partial y^o}{\partial c_1}$, we have $\frac{\partial \Pi_m^*}{\partial c_1} > 0$ if and only if

$$\begin{aligned}
&\sum_{t=l,h} \left(\int_{\underline{a}}^{x_t} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=2}^{n} K_j^*(S_1))^+ - (a + \frac{\epsilon}{b_t} - \sum_{j=1}^{n} K_j^*(S_1))^+ dF(\epsilon) dG(a) \right. \\
&\left. + \int_{x_t}^{\bar{a}} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=2}^{n} K_j^*(S_2))^+ - (a + \frac{\epsilon}{b_t} - \sum_{j=1}^{n} K_j^*(S_2))^+ dF(\epsilon) dG(a) \right) \\
&< \left(-\frac{\partial y^o}{\partial c_1} \right) \sum_{i=1}^{n} \tau(\sum_{j=i}^{n} K_j^*(S_1), \sum_{j=i}^{n} K_j^*(S_2))
\end{aligned} \tag{C.3}$$

Analysis on the Special Case of Uniformly Distributed a and ϵ at $\phi = 1/2$.

As in the proof of Proposition 2, we impose regularity conditions such that for $i = 1, 2, ..., n$, $\sum_{j=i}^{n} K_j^*(m)$ are such that $(\sum_{j=i}^{n} K_j^*(m) - a)b$ is within the support of ϵ for any $(a, b) \in S$ and $m = S_1, S_2$. Note that for $n = 2$, this reduces exactly to the conditions imposed for Proposition 2(i). That is, $\Psi_0(\underline{a} + \frac{\sigma}{b_h}) \leq \frac{r_1}{w - c_1}$ and $\min_{i=1,2} \Psi_i(\bar{a} - \frac{\sigma}{b_h}; x_l, x_h) \geq \frac{r_n - r_{n-1}}{c_{n-1} - c_n}$. Consequently, we can invoke $\bar{F}(x) = \frac{\sigma - x}{2\sigma}$ in the optimality conditions, and obtain $\sum_{j=i}^{n} K_j^*(S_1) = \frac{\bar{a} + \underline{a}}{2} - \frac{\sigma(2z_i - 1)}{b_l}$ and $\sum_{j=i}^{n} K_j^*(S_2) = \frac{\bar{a} + \underline{a}}{2} - \frac{\sigma(2z_i - 1)}{b_h}$ for $i = 2, 3, ..., n$.

Furthermore, in the special case of $\phi = \frac{1}{2}$, $K_T^*(S_1) = K_T^*(S_2) = \frac{\bar{a} + \underline{a}}{2}$. Thus, $K_1^*(S_1) = \frac{\sigma(2z_2 - 1)}{b_l}$ and $K_1^*(S_2) = \frac{\sigma(2z_2 - 1)}{b_h}$.

For the special case of $\phi = \frac{r_1}{w - c_1} = \frac{1}{2}$, we derive more explicit conditions under which Π_m^* is increasing in r_1 and c_1. First, note that x_l and x_h are not continuous at $\phi = \frac{1}{2}$. We focus on the right derivatives of Π_m^*, denoted by $\frac{\partial \Pi_m^*}{\partial r_1^+}$ and $\frac{\partial \Pi_m^*}{\partial c_1^+}$. That is, we consider a small increase in r_1 and c_1 such that ϕ approaches $\frac{1}{2}$ from the left (i.e., in the domain of $\phi \leq \frac{1}{2}$). When $\phi = 1/2$, $x_l = \bar{a}$ and $x_h = \underline{a}$.

Let $z_i = \frac{r_i - r_{i-1}}{c_{i-1} - c_i}$ for $i = 1, ..., n$, where $r_0 = 0$ and $c_0 = w$. We must have $z_n > z_{n-1} > ...z_2 > z_1 = \phi$, given that the n sources are used in the optimal portfolio, i.e., $K_i^*(m) > 0$ for all $i = 1, 2, ..., n$ and $m \in \{S_1, S_2\}$. (See Lemma A.1 for the characterization of the optimal portfolio.)

Substituting these expressions, we obtain

$$\tau(\sum_{j=i}^{n} K_j^*(S_1), \sum_{j=i}^{n} K_j^*(S_2)) = \frac{(b_h^2 - b_l^2)(2z_i - 1)[(b_h - b_l)(2z_i - 1)\sigma - (\bar{a} - \underline{a}) b_h b_l]}{4(\bar{a} - \underline{a}) b_h^2 b_l^2} \tag{C.4}$$

where $z_i = \frac{r_i - r_{i-1}}{c_{i-1} - c_i}$ for $i = 2, ..., n$.

The derivative with respect to r_1. Note that $\frac{\partial y^o}{\partial r_1} = \frac{\partial y^o}{\partial \phi} \frac{\partial \phi}{\partial r_1}$ where $\frac{\partial \phi}{\partial r_1} = \frac{1}{w - c_1}$ and, as derived in the proof of Corollary 1, for the case of $\phi \leq \frac{1}{2}$,

$$\frac{\partial y^o}{\partial \phi} = -\frac{V - V^2(1-2\phi)/\sqrt{(\bar{a}-\underline{a})^2B^4 + V^2}}{B^2}$$

where $B = (b_l + b_h)/2$ and $V = \sigma(b_h - b_l)$. Hence, at $\phi = \frac{1}{2}$, $\frac{\partial y^o}{\partial r_1^+} = -\frac{4(b_h-b_l)\sigma}{(w-c_1)(b_h+b_l)^2}$.

Together with the expression (C.4) and condition (C.2), it follows that $\frac{\partial \Pi_m^*}{\partial r_1^+} > 0$ if and only if

$$\bar{a} - \underline{a} < \frac{(b_h - b_l)^3\sigma \sum_{i=2}^{n}(2z_i - 1)^2}{b_h b_l\left((b_h-b_l)^2\sum_{i=2}^{n}(2z_i-1) + (2z_2-1)(b_h+b_l)^2(w-c_1)\right)}. \quad \text{(C.5)}$$

For $n = 2$, the above reduces to

$$\bar{a} - \underline{a} < \frac{(b_h - b_l)^3\sigma(2z_2 - 1)}{b_h b_l\left((b_h-b_l)^2 + (b_h+b_l)^2(w-c_1)\right)}. \quad \text{(C.6)}$$

The derivative with respect to c_1. Under the special case and $\phi = \frac{1}{2}$, the left-hand side of (C.3) reduces to $\frac{(bh+bl)(4z^2-1)}{4b_hb_l}$. Similarly, $\frac{\partial y^o}{\partial c_1} = \frac{\partial y^o}{\partial \phi}\frac{\partial \phi}{\partial c_1}$. With $\frac{\partial \phi}{\partial c_1} = \frac{r_1}{(w-c_1)^2}$ and the expression of $\frac{\partial y^o}{\partial \phi}$ derived earlier, we have $\frac{\partial y^o}{\partial c_1^+} = -\frac{2(b_h-b_l)\sigma}{(w-c_1)(b_h+b_l)^2}$ at $\phi = \frac{1}{2}$.

With (C.4) and the expressions of $K_2^*(S_1)$ and $K_2^*(S_2)$, we can simplify (C.3) as

$$\bar{a} - \underline{a} < \frac{2(b_h - b_l)^3\sigma \sum_{i=2}^{n}(2z_i - 1)^2}{b_h b_l(2(b_h-b_l)^2\sum_{i=2}^{n}(2z_i-1) + (b_h+b_l)^2(4z_2^2-1)(w-c_1))} \quad \text{(C.7)}$$

For $n = 2$, the above reduces to

$$\bar{a} - \underline{a} < \frac{2(b_h - b_l)^3(2z_2 - 1)\sigma}{b_h b_l(2(b_h-b_l)^2 + (b_h+b_l)^2(2z_2+1)(w-c_1))}. \quad \text{(C.8)}$$

Because Π_m^* and its derivatives are continuous for $\phi \in (0, \frac{1}{2}]$, $\frac{\partial \Pi_m^*}{\partial r_1^+} > 0$ and $\frac{\partial \Pi_m^*}{\partial c_1^+} > 0$ imply that there is a neighbor of $\phi \in [\frac{1}{2} - \delta, \frac{1}{2}]$ in which Π_m^* is increasing. By (C.6) and (C.8), a sufficient condition for $\frac{\partial \Pi_m^*}{\partial r_1^+} > 0$ and $\frac{\partial \Pi_m^*}{\partial c_1^+} > 0$ is that $\bar{a} - \underline{a} < \Gamma$ where

$$\Gamma = \min\left(\frac{(b_h-b_l)^3(2z_2-1)\sigma}{b_hb_l\left((b_h-b_l)^2 + (b_h+b_l)^2(w-c_1)\right)}, \frac{2(b_h-b_l)^3(2z_2-1)\sigma}{b_hb_l(2(b_h-b_l)^2 + (b_h+b_l)^2(2z_2+1)(w-c_1))}\right).$$

Since $z_2 > \phi$, $2z_2 - 1 > 0$ at $\phi = \frac{1}{2}$. Hence, Γ is increasing in σ. □

Proof of Lemma 3. Lemma 3 is an abridged version of Lemma A.1 in Appendix A. In what follows, we prove the optimality of Algorithm A.1 as stated in Lemma A.1, which implies the properties summarized in Lemma 3. Note that source κ is denoted by i_A in Algorithm A.1 as well as the following proof.

For the execution problem (A.2), a greedy solution is clearly optimal: We should execute the capacity with the lowest c_i first until the capacity reserved is saturated or demand is fully satisfied. Recall that $c_n < c_{n-1} < ... < c_1$. Therefore, given any realized demand D and capacity reservation $\mathbf{K}$, the optimal solution to (A.2) is given by the following:

$$x_n^* = \min(K_n, D),$$

$$x_{n-1}^* = \min(K_{n-1} + K_n, D) - \min(K_n, D),$$

$$x_j^* = \min\left(\sum_{i=j}^{n} K_i, D\right) - \min\left(\sum_{i=j+1}^{n} K_i, D\right) \text{ for } j = n-2, n-3, ..., 1$$

Substituting the optimal x_i^*'s into the first-stage reservation problem, we can simplify the manufacturer's objective function (A.1) as

$$\Pi_m(\mathbf{K}|m) = (c_0 - c_n)\mathbb{E}[D|m] - \sum_{i=1}^{n}(c_{i-1} - c_i)\mathbb{E}[(D - \sum_{j=i}^{n} K_j)^+|m] - \sum_{i=1}^{n} r_i K_i \tag{C.9}$$

where we define $c_0 = w$. It is easy to see that (C.9) is concave in $\mathbf{K}$, and so it suffices to analyze the first-order optimality condition. Taking the derivative with respect to each K_i, we have

$$\begin{aligned} \frac{\partial \Pi_m}{\partial K_1} &= (c_0 - c_1)\bar{F}_D(\sum_{j=1}^{n} K_j) - r_1, \\ \frac{\partial \Pi_m}{\partial K_i} &= \sum_{t=1}^{i}(c_{t-1} - c_t)\bar{F}_D(\sum_{j=t}^{n} K_j) - r_i \text{ for } i = 2, 3, ..., n. \end{aligned} \tag{C.10}$$

In the optimal solution, $\frac{\partial \Pi_m}{\partial K_i} = 0$ if $K_i^* > 0$ and $\frac{\partial \Pi_m}{\partial K_i} \leq 0$ if $K_i^* = 0$.

Let I^* denote the set of capacity sources being used in the optimal solution, i.e., $I^* = \{i | K_i^* > 0, i = 1, 2, ..., n\}$. We first establish some properties of the optimal solution.

(1) The smallest index in I^* is given by $i_A = \arg\min_i \frac{r_i}{w - c_i}$; moreover, the total capacity is such that $\bar{F}_D(\sum_{j=1}^{n} K_j^*) = \frac{r_{i_A}}{w - c_{i_A}}$.

To show property (1), first note that if capacity source i is one with the smallest i used in the optimal solution, i.e., $K_i^* > 0$ and $K_l^* = 0$ for all $l < i$, then the optimality conditions imply

$$\frac{\partial \Pi_m}{\partial K_i} = (w - c_i)\bar{F}_D(\sum_{j=i}^{n} K_j^*) - r_i = 0;$$
$$\frac{\partial \Pi_m}{\partial K_l} = (w - c_l)\bar{F}_D(\sum_{j=i}^{n} K_j^*) - r_l \leq 0 \text{ for } l = 1, ..., i' - 1.$$

Hence, we must have $\bar{F}_D(\sum_{j=i}^{n} K_j^*) = \frac{r_i}{w-c_i}$ and $\frac{r_i}{w-c_i} \leq \frac{r_l}{w-c_l}$ for all $l < i$.

On the other hand, suppose that there exits some $l > i$ such that $\frac{r_i}{w-c_i} > \frac{r_l}{w-c_l}$. It follows that $\frac{\partial \Pi_m}{\partial K_l} = (w - c_l)\bar{F}_D(\sum_{j=l}^{n} K_j^*) - r_l \geq (w - c_l)\bar{F}_D(\sum_{j=i}^{n} K_j^*) - r_l = (w - c_l)\frac{r_i}{w-c_i} - r_l > 0$. (The first inequality holds as $\sum_{j=l}^{n} K_j^* \leq \sum_{j=i}^{n} K_j^*$; the second follows by the optimality condition $\frac{\partial \Pi_m}{\partial K_i} = 0$; the last inequality is due to the contradictory hypothesis.) The above implies the optimality condition with respect to K_l can never hold, thus leading to a contradiction. Therefore, i is the index with the smallest $\frac{r_i}{w-c_i}$.

(2) If i and i' are consecutive indexes in I^* where $i < i'$, i.e., $K_i^* > 0, K_{i'}^* > 0$ and $K_l^* = 0$ for all $l \in (i, i')$, then $r_{i'} + c_{i'} < r_i + c_i$, and $\frac{r_{i'} - r_i}{c_i - c_{i'}} \geq \frac{r_{i'} - r_l}{c_l - c_{i'}}$ for all $l \in (i, i')$; moreover, $\bar{F}_D(\sum_{j=i'}^{n} K_j^*) = \frac{r_{i'} - r_i}{c_i - c_{i'}}$.

If in the optimal solution, for some $i < i'$, $K_i^* > 0, K_{i'}^* > 0$ and $K_l^* = 0$ for all $l \in (i, i')$, then the optimality conditions imply

$$\frac{\partial \Pi_m}{\partial K_i} = \sum_{t=1}^{i}(c_{t-1} - c_t)\bar{F}_D(\sum_{j=t}^{n} K_j^*) - r_i = 0;$$
$$\frac{\partial \Pi_m}{\partial K_l} = \sum_{t=1}^{l}(c_{t-1} - c_t)\bar{F}_D(\sum_{j=t}^{n} K_j^*) - r_l \leq 0 \text{ for } l = i + 1, ..., i' - 1;$$
$$\frac{\partial \Pi_m}{\partial K_{i'}} = \sum_{t=1}^{i'}(c_{t-1} - c_t)\bar{F}_D(\sum_{j=t}^{n} K_j^*) - r_{i'} = 0.$$

Since $K_l^* = 0$ for all $l = i + 1, ..., i' - 1$, we can derive the following relations:

$$\frac{\partial \Pi_m}{\partial K_{i'}} - \frac{\partial \Pi_m}{\partial K_i} = (c_i - c_{i'})\bar{F}_D(\sum_{j=i'}^{n} K_j^*) - (r_{i'} - r_i) = 0$$

and

$$\frac{\partial \Pi_m}{\partial K_{i'}} - \frac{\partial \Pi_m}{\partial K_l} = (c_l - c_{i'})\bar{F}_D(\sum_{j=i'}^{n} K_j^*) - (r_{i'} - r_l) \geq 0.$$

Therefore, it must be the case that $\bar{F}_D(\sum_{j=i'}^n K_j^*) = \frac{r_{i'}-r_i}{c_i-c_{i'}}$ and $\frac{r_{i'}-r_i}{c_i-c_{i'}} \geq \frac{r_{i'}-r_l}{c_l-c_{i'}}$ for all $l = i+1, ..., i'-1$. Moreover, for an optimal $\sum_{j=i'}^n K_j^*$ to exist, we must have $\frac{r_{i'}-r_i}{c_i-c_{i'}} < 1$, i.e., $r_{i'} + c_{i'} < r_i + c_i$.

(3) The largest index in I^* is given by $i_B = \arg\min_i r_i + c_i$.

Let i' be the largest index in I^*. Then, we can show that $r_i' + c_i' < r_l + c_l$ for any $l < i'$ using Properties (1) and (2) above. To see this, note that $i_A \leq i'$ where i_A is as defined in Property (1). Property (1) implies that for any $l < i_A$, $\frac{r_{i_A}}{w-c_{i_A}} \leq \frac{r_l}{w-c_l}$, which holds if and only if $\frac{r_{i_A}-r_l}{c_l-c_{i_A}} \leq \frac{r_{i_A}}{w-c_{i_A}}$ (by rearranging the terms). Since $\frac{r_{i_A}}{w-c_{i_A}} < 1$, it follows that $r_{i_A} + c_{i_A} < r_l + c_l$ whenever $l < i_A$. Furthermore, Property (2) implies that for any $i < i'$ in I^*, $r_{i'} + c_{i'} < r_i + c_i$. Moreover, for any l between i and i', we have $1 > \frac{r_{i'}-r_i}{c_i-c_{i'}} \geq \frac{r_{i'}-r_l}{c_l-c_{i'}}$, and so $r_{i'} + c_{i'} < r_l + c_l$. Combining the above yields $r_{i'} + c_{i'} < r_l + c_l$ for all $l < i'$.

Now consider $l > i'$. Because i' is the largest index in I^*, we have $K_{i'}^* > 0$ and $K_l^* = 0$. The optimality conditions then imply $\frac{\partial \Pi_m}{\partial K_l} \leq 0$ and $\frac{\partial \Pi_m}{\partial K_{i'}} = 0$. It follows that

$$\begin{aligned} 0 &\geq \frac{\partial \Pi_m}{\partial K_l} - \frac{\partial \Pi_m}{\partial K_{i'}} \\ &= \sum_{t=i'+1}^{l} (c_{t-1} - c_t)\bar{F}_D(\sum_{j=t}^{n} K_j^*) - (r_l - r_{i'}) \\ &= (c_{i'} - c_l) - (r_l - r_{i'}) \end{aligned}$$

where the second inequality follows because i' is the largest index in I^* and so $K_j^* = 0$ for all $j > i'$. Therefore, $r_{i'} + c_{i'} \leq r_l + c_l$ for all $l > i'$. We thus conclude the proof that $i' = i_B = \arg\min_i r_i + c_i$.

By the above three properties, Algorithm A.1 as described in the lemma will determine a set I^* that satisfies optimality conditions. Given I^*, the optimal capacity levels $\mathbf{K}^*$ can be derived according to Properties (1)(2) above.

$\square$

Proof of Proposition 5. Because (i) the retailer's payoff depends only on the total capacity level $\sum_{i=1}^n K_i$, and (ii) the manufacturer's optimal total capacity is determined by $\bar{F}_D(K_T^*|m) = \frac{r_\kappa}{w-c_\kappa}$, the proof for the two-capacity source setting remains valid for the n-capacity source case while the ratio that determines the total capacity becomes $\frac{r_\kappa}{w-c_\kappa}$. $\square$

Proof of Proposition 6. If the addition of j does not change i_A, then the informative partition will remain the same. As the portfolio is expanded, conditioning on each S_i, the manufacturer will solve the same problem except with an additional source j. Thus, the manufacturer will be (weakly) better off with the expanded portfolio.

If j becomes the critical source after expansion, then the example given in the main paper proves statement (ii). Furthermore, statement (ii) can be implied by Proposition A.1 where we show that the manufacturer's expected profit can be increasing in r_1 or c_1 given the optimal portfolio consisting of $n \geq 2$ sources. This is because one can view the addition of source 1 (i.e., the critical source) as source 1's cost r_1 or c_1 decreasing from a sufficiently high level (such that source 1 should be just eliminated) to an attractive level. Specifically, for any three-source instance in which the manufacturer receives a lower payoff with r_1 than with r_1' where $r_1 < r_1'$, one can construct cost parameters (r_2, c_2) and (r_3, c_3) such that at r_1', the manufacturer is indifferent about whether or not to include source 1 in the optimal portfolio. Thus, expanding a portfolio from $\{2, 3\}$ to $\{1, 2, 3\}$ leads to a lower payoff if the reservation cost of source 1 is r_1. □

Proof of Proposition 7. For any given k, define $\Delta(y) = \Psi_1(k; y) - \Psi_2(k; y)$. First, $\Delta(y)$ is continuous in $y \in [0, \pi]$. Second,

$$\Delta(0) = \Psi_1(k;0) - \Psi_2(k;0) = \frac{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{b_m} \bar{F}((k-a)b) g(a,b) dbda}{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{b_m} g(a,b) dbda} - \frac{\int_{\underline{a}}^{\bar{a}} \int_{b_m}^{\bar{b}} \bar{F}((k-a)b) g(a,b) dbda}{\int_{\underline{a}}^{\bar{a}} \int_{b_m}^{\bar{b}} g(a,b) dbda}$$

and

$$\Delta(\pi) = \Psi_1(k;\pi) - \Psi_2(k;\pi) = \frac{\int_{\underline{a}}^{\bar{a}} \int_{b_m}^{\bar{b}} \bar{F}((k-a)b) g(a,b) dbda}{\int_{\underline{a}}^{\bar{a}} \int_{b_m}^{\bar{b}} g(a,b) dbda} - \frac{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{b_m} \bar{F}((k-a)b) g(a,b) dbda}{\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{b_m} g(a,b) dbda}.$$

So, $\Delta(0) = -\Delta(\pi)$, i.e., $\Delta(0)$ and $\Delta(\pi)$ have opposite signs.

By the Intermediate Value Theorem, there must exist $y \in [0, \pi]$ such that $\Delta(y) = 0$, i.e., $\Psi_1(k; y) = \Psi_2(k; y)$, given any total capacity level $k = K_T$. □

Proof of Proposition 8. Because of the incentive-compatible condition, the total capacity is message-dependent. We show that the total capacity with communication is optimal the manufacturer's problem without communication.

By the optimality condition with the total capacity, $K_T^*(S_i)$ solves $\Psi_i(K_T^*, y) = \phi := \frac{r_\kappa}{w - c_\kappa}$ for both $i = 1, 2$. On the other hand, the optimality condition for the manufacturer's problem absent communication can be written as $\Psi_0(K_T^N) = \phi$ where

$$\Psi_0(k) = \int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{\bar{b}} \bar{F}((k-a)b) g(a,b) dbda.$$

For any $y \in [0, y^\dagger]$, by (B.1) and (B.2), it follows that

$$\left(\int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{(a_m - a)\tan y + b_m} g(a,b) db da\right) \Psi_1(k; y) + \left(\int_{\underline{a}}^{\bar{a}} \int_{(a_m - a)\tan y + b_m}^{\bar{b}} g(a,b) db da\right) \Psi_2(k; y)$$
$$= \int_{\underline{a}}^{\bar{a}} \int_{\underline{b}}^{\bar{b}} \bar{F}((k-a)b) g(a,b) db da = \Psi_0(k).$$

That is, $\Psi_0(k)$ can be written as a convex combination of $\Psi_1(k; y)$ and $\Psi_2(k; y)$. Given the optimality condition of K_T^*, i.e., $\Psi_1(K_T^*; y) = \Psi_2(K_T^*; y) = \phi$, the above reduces to $\phi = \Psi_0(K_T^*)$. That is, the optimality of K_T^* under the informative equilibrium implies that K_T^* is also optimal to the manufacturer's problem without communication.

For the cases of $y \in (y^\dagger, \pi - y^\dagger)$ and $y \in [\pi - y^\dagger, \pi]$, the proof is analogous. We can always write $\Psi_0(k)$ as a convex combination of $\Psi_1(k; y)$ and $\Psi_2(k; y)$, and therefore the optimality of K_T^* with communication implies that it also solves the manufacturer's problem without communication.

The profit implications of communication follows immediately since the retailer's payoff depends only on the total capacity level and the manufacturer is able to make a more informed decision on capacity combination. □

Proof of Proposition 9. By Lemma 2, the optimal total capacities with and without communication are equal, i.e., $K_T^*(S_1) = K_T^*(S_2) = K_T^N$. To determine an informative partition, we need to find (x_l, x_h) that solves equation $\Psi_1(K_T^N; x_l, y_h) = \Psi_2(K_T^N; x_l, y_h)$. As in the proof of Proposition 2, under the condition $\Psi_0(\underline{a} + \frac{\sigma}{b_h}) \leq \frac{r_1}{w - c_1}$, we are able to substitute $G(a) = \frac{a - \underline{a}}{\bar{a} - \underline{a}}$ and $\bar{F}(x) = \frac{\sigma - x}{2\sigma}$, $K_T^N =$ into $\Psi_1(K_T^N; x_l, y_h)$ and $\Psi_2(K_T^N; x_l, y_h)$, where $K_T^N = \frac{\underline{a} + \bar{a}}{2} + \frac{\sigma(1 - 2\phi)}{\rho_h b_h + \rho_l b_l}$ is the optimal total capacity without communication.

Assuming $x_l = \underline{a}, \bar{a}$, respectively, we can find the corresponding x_h satisfying $\Psi_1(K_T^N; x_l, y_h) = \Psi_2(K_T^N; x_l, y_h)$:

$$x_l = \underline{a}, x_h = \bar{a} - (2\phi - 1)v_h$$
$$x_l = \bar{a}, x_h = \underline{a} + (1 - 2\phi)v_h$$

where $v_h = \frac{2\sigma\rho_l(b_h - b_l)}{b_h(\rho_l b_l + \rho_h b_h)}$. Likewise, assuming $x_h = \underline{a}, \bar{a}$, respectively, we can solve for the corresponding x_l as follows:

$$x_l = \bar{a} - (1 - 2\phi)v_l, x_h = \underline{a};$$
$$x_l = \underline{a} + (2\phi - 1)v_l, x_h = \bar{a}.$$

where $v_l = \frac{2\sigma\rho_h(b_h - b_l)}{b_l(\rho_l b_l + \rho_h b_h)}$.

Note that x_l and x_h need to be within $[\underline{a}, \bar{a}]$. Therefore, we impose a condition $\bar{a} - \underline{a} \geq |1 - 2\phi| \max(v_l, v_h)$, and choose the solutions satisfying $\underline{a} \leq x_l, x_h \leq \bar{a}$, depending on whether $\phi \leq \frac{1}{2}$ [case (a)] or $\phi > \frac{1}{2}$ [case (b)].

Next, we characterize the relation between $K_2^*(S_1)$ and $K_2^*(S_2)$. As in the proof of Proposition 2, we evaluate the derivative of $\Delta(k) = \Psi_1(k; x_l, x_h) - \Psi_2(k; x_l, x_h)$. The condition $\min_{i=1,2} \Psi_i(\bar{a} - \frac{\sigma}{b_h}; x_l, x_h) \geq \frac{r_2 - r_1}{c_1 - c_2}$ ensures that $(K_2^*(S_i) - a)b$ falls within the support of ϵ for all $i = 1, 2$ and $(a, b) \in S$. So, we can simplify $\Delta'(k)$ with $\bar{F}(x) = \frac{\sigma - x}{2\sigma}$. It follows that

$$\Delta'(k) = -\frac{(\bar{a} - \underline{a})(b_h - b_l)\rho_h \rho_l (x_h - x_l)}{2\sigma \sum_{t \in l,h} \rho_t (x_t - \underline{a}) \sum_{t \in l,h} \rho_t (\bar{a} - x_t)}.$$

As a result, $\Delta'(k) > 0$ for $x_l > x_h$, and $\Delta'(k) < 0$ for $x_l < x_h$.

Therefore, for case (a), $\Delta'(k) > 0$. Moreover $\Delta(K_T^*) = 0$. For any $k < K_T^*$, $\Delta(k) < 0$, i.e., $\Psi_1(k; x_l, x_h) < \Psi_2(k; x_l, x_h)$. On the other hand, for $i = 1, 2$, the optimality condition of $K_2^*(S_i)$ is given by $\frac{r_2 - r_1}{c_1 - c_2} = \Psi_i(K_2^*(S_i); x_l, x_h)$. Because $\Psi_i(k)$ is decreasing in k, it must be the case that $K_2^*(S_1) < K_2^*(S_2)$. For case (b), $\Delta'(k) < 0$. By a similar argument with the optimality condition, it follows that $K_2^*(S_1) > K_2^*(S_2)$.

Since the total capacity levels are identical for both S_1 and S_2, the relation between $K_1^*(S_1)$ and $K_1^*(S_2)$ follows automatically. □

D Remark on the Regularity Conditions of Proposition 2

For analytical convenience, we impose two regularity conditions (C1)(C2) when proving Proposition 2. These two conditions together ensure that $-\sigma \leq (K_2^*(S_i) - a)b \leq (K_T^* - a)b \leq \sigma$ for all $(a, b) \in S$ and $i = 1, 2$, and we can thereby invoke the expression $\bar{F}(x) = \frac{\sigma - x}{2\sigma}$ without worrying that x goes out of the uniform support $[-\sigma, \sigma]$. In the following, we derive more explicit sufficient conditions for the two regularity conditions to hold under the special case of $\rho_l = \rho_h = \frac{1}{2}$. As shown below, the sufficient conditions require that $\bar{a} - \underline{a}$ is small enough compared to 2σ, i.e., the support of a is narrow enough relative to the support of noise ϵ.

Condition (C1). With the expression of $\Psi_0(k)$, Condition (C1) can be written as

$$\sum_{t=l,h} \rho_t \int_{\underline{a}}^{\bar{a}} \bar{F}((\underline{a} + \frac{\sigma}{b_h} - a)b_t) dG(a) \leq \phi.$$

To simplify the condition, notice the following. First, for all a and b_t where $t = l, h$, $(\underline{a} + \frac{\sigma}{b_h} - a)b_t \leq \frac{b_t \sigma}{b_h} \leq \sigma$. Second, for all a and b_t, $(\underline{a} + \frac{\sigma}{b_h} - a)b_t \geq (\underline{a} + \frac{\sigma}{b_h} - \bar{a})b_t \geq -\sigma$ provided that $\bar{a} - \underline{a} \leq \frac{2\sigma}{b_h}$.

Assuming $\bar{a} - \underline{a} \leq \frac{2\sigma}{b_h}$, we can then invoke $\bar{F}(x) = \frac{\sigma - x}{2\sigma}$ in the above expression since the argument must be within $[-\sigma, \sigma]$. Then, simplifying the integrals leads to

$$\sum_{t=l,h} \rho_t \int_{\underline{a}}^{\bar{a}} \bar{F}((\underline{a} + \frac{\sigma}{b_h} - a)b_t) dG(a) = \frac{(\bar{a} - \underline{a})(\rho_l b_l + \rho_h b_h)}{4\sigma} + \frac{(b_h - b_l)\rho_l}{2b_h}.$$

Rearranging the terms, Condition (C1) can be expressed as $\bar{a} - \underline{a} \leq 2\sigma \left(\frac{1}{b_h} + \frac{2\phi - 1}{\rho_l b_l + \rho_h b_h} \right)$. Together with the assumption made earlier, we conclude that Condition (C1) is satisfied if $\bar{a} - \underline{a} \leq 2\sigma \min \left(\frac{1}{b_h}, \frac{1}{b_h} + \frac{2\phi - 1}{\rho_l b_l + \rho_h b_h} \right)$.

Condition (C2). It is more involved to analyze Condition (C2), and so we focus on the special case of $\rho_l = \rho_h = \frac{1}{2}$. With $G(a) \sim U[\underline{a}, \bar{a}]$, Condition (C2) can be written as follows. for $i = 1$,

$$\Psi_1(\bar{a} - \frac{\sigma}{b_h}; x_l, x_h) = \frac{1}{2(\bar{a} - \underline{a})} \left(\int_{\underline{a}}^{x_l} \bar{F}((\bar{a} - \frac{\sigma}{b_h} - a)b_l) da + \int_{\underline{a}}^{x_h} \bar{F}((\bar{a} - \frac{\sigma}{b_h} - a)b_h) da \right)$$

and

$$\Psi_2(\bar{a} - \frac{\sigma}{b_h}; x_l, x_h) = \frac{1}{2(\bar{a} - \underline{a})} \left(\int_{x_l}^{\bar{a}} \bar{F}((\bar{a} - \frac{\sigma}{b_h} - a)b_l) da + \int_{x_h}^{\bar{a}} \bar{F}((\bar{a} - \frac{\sigma}{b_h} - a)b_h) da \right).$$

When evaluating the integrals above, in general, one has to consider the three cases of $\bar{F}(x)$, that is, $\bar{F}(x) = \begin{cases} 1 \text{ if } x \leq -\sigma, \\ \frac{\sigma - x}{2\sigma} \text{ if } -\sigma < x \leq \sigma, \\ 0 \text{ if } x > \sigma. \end{cases}$

For all $t = l, h$, on the one hand, note that $(\bar{a} - \frac{\sigma}{b_h} - a)b_t \geq -\frac{\sigma b_t}{b_h} \geq -\sigma$; on the other hand, $(\bar{a} - \frac{\sigma}{b_h} - a)b_t \leq (\bar{a} - \frac{\sigma}{b_h} - \underline{a})b_t$, which will be lower than or equal to σ under a sufficient condition $(\bar{a} - \underline{a}) \leq \frac{2\sigma}{b_h}$.

Hence, assuming $\bar{a} - \underline{a} \leq \frac{2\sigma}{b_h}$, we can invoke $\bar{F}(x) = \frac{\sigma - x}{2\sigma}$ in the expressions of Ψ_1 and Ψ_2. It follows that

$$\Psi_1(\bar{a} - \frac{\sigma}{b_h}; a_m - y, a_m + y) = \frac{4\sigma(\bar{a} - \underline{a})(3b_h + b_l) - 3b_h(b_h + b_l)(\bar{a} - \underline{a})^2 + 4W(y)}{16\sigma b_h(\bar{a} - \underline{a})}$$

where $W(y) = b_h(b_h + b_l)y^2 - (b_h - b_l)(b_h(\bar{a} - \underline{a}) - 2\sigma)y$. An exact analysis on Ψ_1 is intractable since in equilibrium the value of y is determine by the involved expression (6). Hence, we derive a lower bound of Ψ_1 by minimizing $W(y)$ over y. It is easy to see that $W(y)$ is a quadratic function of y with an unconstrained minimizer $y^+ = -\frac{(b_h-b_l)(2\sigma-b_h(\bar{a}-\underline{a}))}{2b_h(b_h+b_l)}$. Note that $y^+ \le 0$ under the assumption of $\bar{a}-\underline{a} \le \frac{2\sigma}{b_h}$. Because $y \in [-\frac{\bar{a}-\underline{a}}{2}, \frac{\bar{a}-\underline{a}}{2}]$ and $W(y)$ is symmetric about y^+, it follows that $W(y)$ is minimized at the boundary $y = -\frac{\bar{a}-\underline{a}}{2}$ if $-\frac{\bar{a}-\underline{a}}{2} \ge y^+$. By rearranging the terms, $-\frac{\bar{a}-\underline{a}}{2} \ge y^+$ is equivalent to $\bar{a}-\underline{a} \le \frac{2\sigma(b_h-b_l)}{3b_h^2+b_hb_l}$.

Therefore, if $\bar{a}-\underline{a} \le \min\{\frac{2\sigma}{b_h}, \frac{2\sigma(b_h-b_l)}{3b_h^2+b_hb_l}\}$, setting $y = -\frac{\bar{a}-\underline{a}}{2}$, we can obtain a lower bound of $\Psi_1(\bar{a}-\frac{\sigma}{b_h}; a_m-y, a_m+y)$ as follows

$$\Psi_1(\bar{a}-\frac{\sigma}{b_h}; a_m-y, a_m+y) \ge \Psi_1(\bar{a}-\frac{\sigma}{b_h}; \bar{a}, \underline{a}) = \frac{-b_hb_l(\bar{a}-\underline{a})+2\sigma(b_h+b_l)}{4\sigma b_h}.$$

As a result, a sufficient condition for $\Psi_1(\bar{a}-\frac{\sigma}{b_h}; a_m-y, a_m+y) \ge \frac{r_2-r_1}{c_1-c_2}$ can be written as $\bar{a}-\underline{a} \le \frac{2\sigma(b_h(1-2z)+b_l)}{b_hb_l}$ where $z = \frac{r_2-r_1}{c_1-c_2}$.

Following a similar approach, under the assumption of $\bar{a}-\underline{a} \le \frac{2\sigma}{b_h}$, we can write $\Psi_2(\bar{a}-\frac{\sigma}{b_h}; a_m - y, a_m+y)$ as

$$\Psi_2(\bar{a}-\frac{\sigma}{b_h}; a_m-y, a_m+y) = \frac{4\sigma(\bar{a}-\underline{a})(3b_h+b_l)-(\bar{a}-\underline{a})^2b_h(b_h+b_l)-4W(y)}{16\sigma b_h(\bar{a}-\underline{a})}$$

where $W(y) = b_h(b_h+b_l)y^2+(b_h-b_l)(2\sigma-(\bar{a}-\underline{a})b_h)y$. To derive a lower bound of Ψ_2, we notice that $W(y)$ is a quadratic convex function and symmetric about $y^+ = -\frac{(b_h-b_l)(2\sigma-(\bar{a}-\underline{a})b_h)}{2b_h(b_h+b_l)}$. Note that $y^+ \le 0$ by the assumption of $\bar{a}-\underline{a} \le \frac{2\sigma}{b_h}$. Therefore, $W(y)$ is maximized at $y = \frac{\bar{a}-\underline{a}}{2}$, and so a lower bound of Ψ_2 can be obtained by setting $y = \frac{\bar{a}-\underline{a}}{2}$:

$$\Psi_2(\bar{a}-\frac{\sigma}{b_h}; a_m-y, a_m+y) \ge \Psi_2(\bar{a}-\frac{\sigma}{b_h}; \underline{a}, \bar{a}) = \frac{-b_hb_l(\bar{a}-\underline{a})+2\sigma(b_h+b_l)}{4\sigma b_h}.$$

Thus, a sufficient condition for $\Psi_2(\bar{a}-\frac{\sigma}{b_h}; a_m-y, a_m+y) \ge z$ is $\bar{a}-\underline{a} \le \frac{2\sigma(b_h(1-2z)+b_l)}{b_hb_l}$ where $z = \frac{r_2-r_1}{c_1-c_2}$.

Putting together the conditions derived above, we conclude that if $\bar{a}-\underline{a} \le 2\sigma \min\left(\frac{1}{b_h}, \frac{(b_h(1-2z)+b_l)}{b_hb_l}, \frac{(b_h-b_l)}{3b_h^2+b_hb_l}\right)$, Condition (C2) is satisfied.

In summary, for each regularity condition to hold, we have derived a sufficient condition which can be written as $\bar{a}-\underline{a}$ being not greater than 2σ times a given factor. In other words, conditions (C1) and (C2) hold if the support of ϵ, i.e., $[-\sigma,\sigma]$, is sufficiently wide relative to the support of a, i.e., $[\underline{a}, \bar{a}]$.

E Additional Numerical Results for the Base Model

E.1 Additional Numerical Study on the Value of Communication

As a supplement to the numerical examples in §4 the main paper, using a larger numerical study, we further explore the value of communication to the manufacturer, defined as the percentage improvement in the manufacturer's expected profit due to informative communication, i.e., $(\Pi_m^* - \Pi_m^N)/\Pi_m^N$. We consider the case where the noise term ϵ follows a truncated normal distribution over $[-500, 500]$ with mean 0 and standard deviation denoted by std.

Table 2 reports three representative instances in which $\phi = 0.4, 0.5$ and 0.67, respectively. The observations are similar to the examples discussed in the main paper.

Table 2: Numerical examples of influential communication with truncated normal noise $\epsilon \sim N(0, 300^2)$ truncated in $[-500, 500]$ (Other parameters: $a \sim U[500, 600]$, $(r_1, c_1) = (1, 4.5)$, $(r_2, c_2) = (5, 0)$, $b_l = 1, b_h = 10$ and $\rho_l = \rho_h = 0.5$)

		No Communication		**With Communication**			
w	ϕ	K_1^N, K_2^N	Π_M^N	$K_1^*(S_1), K_2^*(S_1)$	$K_1^*(S_2), K_2^*(S_2)$	Π_M^*	$\frac{\Pi_m^* - \Pi_m^N}{\Pi_m^N}$ (%)
7	0.40	222.6, 345.6	794.7	319.2, 249.0	63.5, 504.7	828.9	4.30
6.5	0.50	204.4, 345.6	546.3	314.6, 235.4	47.7, 502.3	594.8	8.88
6	0.67	172.6, 345.5	304.0	50.0, 468.1	232.5, 285.6	314.9	3.60

Moreover, we maintain the capacity costs (r_i, c_i)'s and the prior distribution of a as reported in Table 2. We compute the manufacturer's expected profits with and without communication as the wholesale price w varies from 5.7 to 9.5 in increments of 0.2. Figure E.1 depicts the value of communication to the manufacturer as a function of w; in Panel (a), we vary the std of ϵ with value b_h/b_l fixed as 10; in Panel (b), we change the ratio b_h/b_l (while $b_l = 1$) with $std = 300$.

Several observations are worth noting. First, and as discussed above, the value of communication (to the manufacturer) is highest at $w = 6.5$ because the critical fractile equals 0.5 at this price and full revelation of the forecast accuracy occurs. The value of communication reduces as the price (and hence the critical fractile) diverges in either direction from this maximizing point because less information revelation occurs (though some information is still revealed). Second, Figure E.1 shows that, when w is not far away from 6.5 such that forecast accuracy is fully or largely revealed, the value of communication increases in the forecast noise uncertainty [Panel (a)] and the information asymmetry as measured by b_h/b_l [Panel (b)]. Intuitively, information about forecast accuracy b is more valuable to the manufacturer as the variability of ϵ or the ratio b_h/b_l is larger. However,

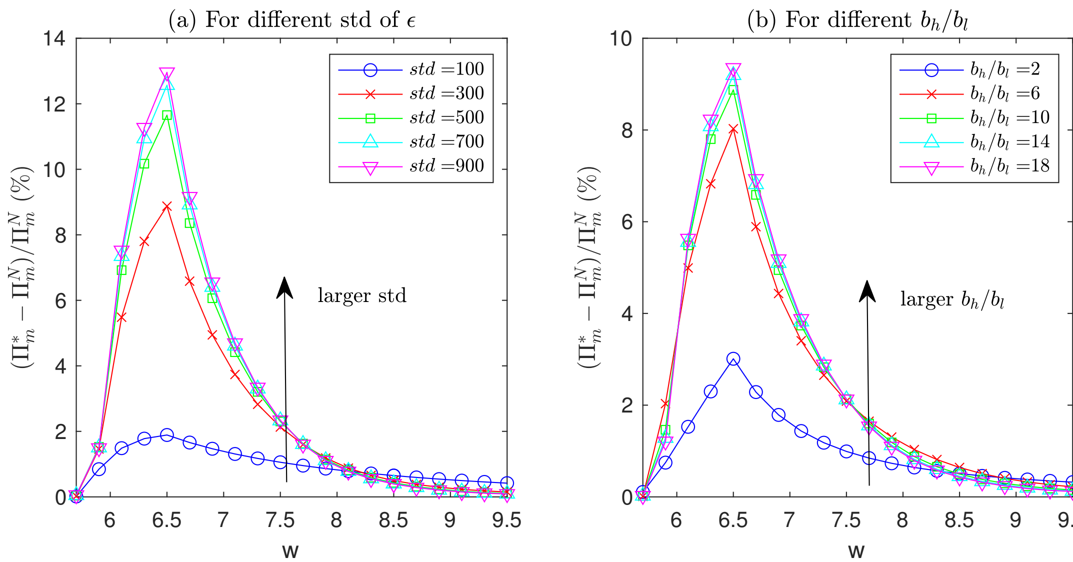

Figure E.1: Value of communication (to the manufacturer) as a function of w. $[(r_1, c_1) = (1, 4.5)$, $(r_2, c_2) = (5, 0)$, $\epsilon \sim$ truncated Normal$(0, std^2)$, $a \sim U[500, 600]$, $b_l = 1$; in panel (a), $b_h = 10$ and std is varied; in panel (b) $std = 300$ and b_h is varied.]

when w is lower than 6 or larger than 8 in the example of Figure E.1, the effects of greater noise uncertainty and information asymmetry can be (slightly) reversed. The reason is that a greater noise uncertainty or information asymmetry, although increasing the value of revealing forecast accuracy, also makes it harder for the retailer to truthfully communicate forecast accuracy in equilibrium. Specifically, we observed that as std or b_h/b_l increases, the informative partition changes with x_l and x_h moving towards the mid-point a_m, i.e., less forecast-accuracy information is revealed. The reduced information revelation caused by a larger std or b_h/b_l can be a dominating force (but its impact appears limited), when w is far away from the value-maximizing point, or equivalently, ϕ is far from 0.5.

E.2 Numerical Study with Asymmetrically Distributed ϵ

In the main body of the paper, we focus on the cases in which demand noise ϵ follows a symmetric distribution. As a result, full revelation of the accuracy information is achieved when $\phi = 0.5$. This specific value 0.5 can be different if the distribution of ϵ is asymmetric. In this appendix, we examine the cases with asymmetrically distributed ϵ.

Triangular Distribution. We consider ϵ following a class of triangular distributions with mode t_m and support $[-\frac{t_m}{2} - \frac{3a_0}{4}, -\frac{t_m}{2} + \frac{3a_0}{4}]$. This specification ensures that (1) the mean equals zero for any given t_m and (2) the support is lower bounded by $-a_0$ such that the demand will not be negative. The distribution is right-skewed (resp. left-skewed) if $t_m < 0$ (resp. $t_m > 0$). Note that a distribution is said to be left-skewed if it is longer on the left side of its peak than on its right, and vice versa. Setting $t_m \in \{-250, -175, 0, 175, 250\}$, we test five triangular distributions, $Tri(-250, -250, 500)$, $Tri(-287.5, -175, 462.5)$, $Tri(-375, 0, 375)$, $Tri(-462.5, 175, 287.5)$, $Tri(-500, 250, 250)$, varied from right-skewed to left-skewed.

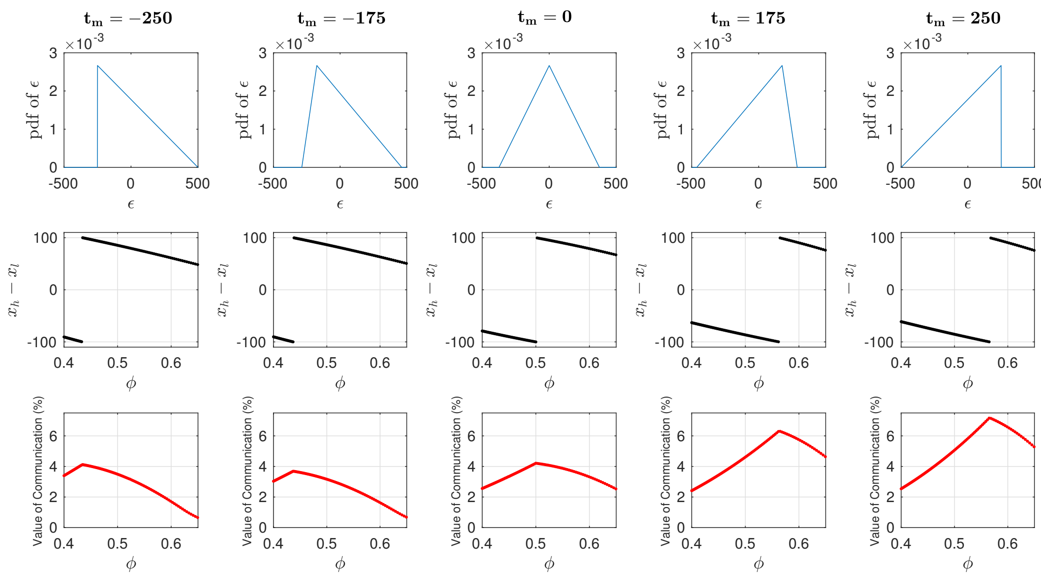

Figure E.2: Effect of distribution skewness of ϵ on informative partitions and value of communication: Triangular Distributions

As reported in Figure E.2, consistent with our observations with symmetric distributions, the value of communication is the highest when $|x_h - x_l|$ is maximized, i.e., accuracy information is fully revealed; see the second and third rows of Figure E.2. However, with asymmetric distributions, the maximum $|x_h - x_l|$ occurs not necessarily at $\phi = 0.5$. As can be seen in the first row of Figure E.2, for right-skewed distributions ($t_m < 0$), $|x_h - x_l|$ is maximized at some $\phi < 0.5$, whereas for left-skewed distributions ($t > 0$), $|x_h - x_l|$ is maximized at some $\phi > 0.5$.

Truncated Skewed Normal Distribution. A skewed Normal distribution has a probability density function $f(x) = 2\varphi(\frac{x-\mu}{\sigma})\Phi(\alpha\frac{x-\mu}{\sigma})$ where φ and Φ represent the pdf and cdf of a standard Normal distribution $N(0,1)$. The distribution $f(x)$ is right-skewed (resp. left-skewed) if $\alpha > 0$ (resp. $\alpha < 0$) and has a mean value of $\xi + \sigma\frac{\alpha}{\sqrt{1+\alpha^2}}\sqrt{\frac{2}{\pi}}$. To ensure that ϵ has a zero mean and that demand is nonnegative, in the experiments, we vary $\alpha \in \{-10, -5, -11, 5, 10\}$ with $\xi = -\sigma\frac{\alpha}{\sqrt{1+\alpha^2}}\sqrt{\frac{2}{\pi}}$, and truncate the distribution within support $[-a_0, a_0]$. We set $\sigma = 500$.

As reported in Figure E.3, we obtain similar observations as in the case of triangular distribu-

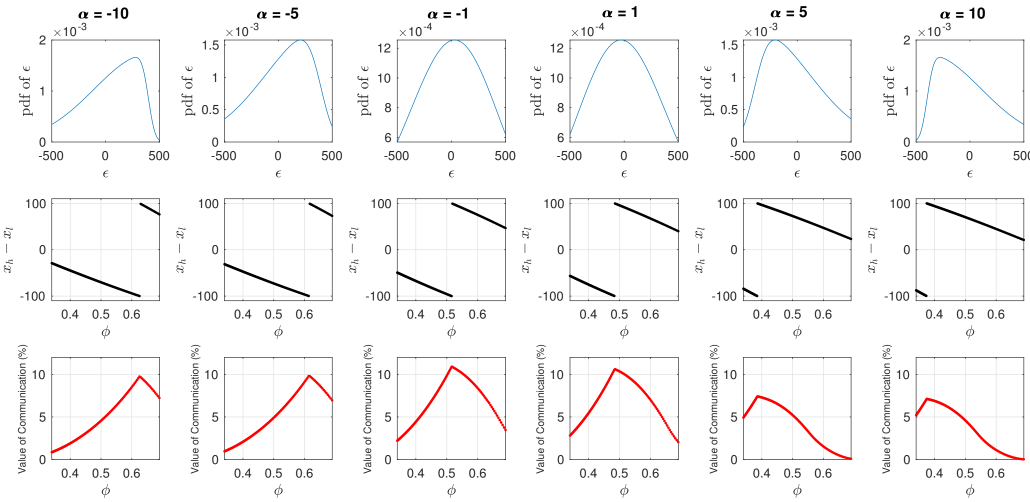

Figure E.3: Effect of distribution skewness of ϵ on informative partitions and value of communication: Truncated Skewed Normal

tions. The full revelation of accuracy information and the maximum value of communication are achieved at some $\phi > 0.5$ (resp. < 0.5) when the distribution is left-skewed (resp. right-skewed).

In summary, we find that *(i) for asymmetrically distributed ϵ, the value of communication is still maximized when forecast accuracy information is fully revealed, and (ii) this happens at some $\phi < 0.5$ for right-skewed distributions but at some $\phi > 0.5$ for left-skewed distributions.* The impact of distribution skewness can be explained intuitively using the newsvendor formula. For a given mean and standard deviation (SD), the order quantity under normal demand is equal to $Mean + zSD$. The SD does not influence the optimal order quantity if z-score equals zero. In such a case, the reported forecast accuracy does not impact the manufacturer's total capacity, thus resulting in full revelation of accuracy information. For a standard normal distribution, a z-score of zero corresponds to an understocking probability of 0.5, and so the full information revelation occurs at $\phi = 0.5$. (Recall that by the optimality condition of our problem, the optimal total capacity K_T^* is such that the understocking probability $\Pr(D > K_T^*) = \phi$.) As the distribution becomes right-skewed, the understocking probability corresponding to $z = 0$ is less than half, and so the full information revelation occurs when $\phi < 0.5$. Similarly, for left-skewed distributions, full information revelation occurs when $\phi > 0.5$.

F Numerical Analysis on Profit-Maximizing Partitions

To determine the informative partition that maximizes the manufacturer's expected profit, one needs to solve the following optimization problem:

$$\max_{(x_l,x_h)\in[\underline{a},\bar{a}]^2} \Pi_m^*(x_l, x_h)$$

subject to

$$\Psi_1(K_T^*; x_l, x_h) = \Psi_2(K_T^*; x_l, x_h) = \min_{1\le i\le n} \frac{r_i}{w - c_i}. \tag{F.1}$$

The objective function $\Pi_m^*(x_l, x_h)$ is the manufacturer's expected profit under any partition parameterized by (x_l, x_h). In the constraint, functions Ψ_1 and Ψ_2 are defined in (4) of the main paper, which are generally nonlinear in (x_l, x_h). As a result, the above problem is a nonconvex optimization. We therefore resort to an exhaustive search to determine the profit-maximizing partition (x_l^*, x_h^*). Specifically, we discretize $[\underline{a}, \bar{a}]$ into a set of N points, denoted by a_i $(i = 1, 2, ..., N)$. For each a_i, we fix $x_l = a_i$ and numerically solve (F.1) to obtain the corresponding $x_h(a_i)$, and then evaluate $\Pi_m^*(a_i, x_h(a_i))$. Skip a_i if no solution $x_h(a_i)$ is found. Lastly, choose the optimal (x_l^*, x_h^*) as the one resulting in the maximum $\Pi_m^*(a_i, x_h(a_i))$.

Consider the following set of parameters: $a \sim U[500, 510]$, $\epsilon \sim$Truncated Normal$(0, 1000^2)$ over $[-500, 500]$, $\rho_h = \rho_l = 0.5$, $b_h = 20$ and $b_l = 1$. There are two capacity sources available with $(r_1, c_1) = (1, 4.5)$ and $(r_2, c_2) = (5, 0)$. We vary the value of w such that $\phi = \frac{r_1}{w-c_1}$ ranges from 0.75 to 0.25 with decrement of 0.05. With the exhaustive search approach described above, we numerically compute the profit-maximizing partition and the resulting equilibrium outcome for each ϕ.[20] Results are reported in Table 3.

For $\phi = 0.5$, the symmetric partition $(x_l^*, x_h^*) = (510, 500)$ also maximizes the manufacturer's profit.[21] For $\phi = 0.45$ and 0.55, the profit-maximizing partition follows the asymmetric form characterized in Proposition 9 in which x_l^* equals endpoint $\underline{a} = 500$ or $\bar{a} = 510$. For other instances, the profit-maximizing partitions are in between, i.e., neither symmetric nor with x_l^* or x_h^* at an endpoint. Nevertheless, all our results about the equilibrium properties continue to hold under the profit-maximizing partition. We have $x_l^* < x_h^*$ ("low-average, high-accuracy" vs "high-average,

[20]Specifically, we discretized $[\underline{a}, \bar{a}]$ into $N = 1001$ points. With x_l being fixed at each of the N points, we solved equation (F.1) to obtain a pair of $(x_l, \hat{x}_h(x_l))$, if any, that constitutes an informative partition. Finally, we selected the pair $(x_l, \hat{x}_h(x_l))$ leading to the maximum manufacturer profit.

[21]Our numerical results suggest that the symmetric partition is also the unique partition that supports an informative equilibrium for the case of $\phi = 0.5$.

Table 3: Equilibrium outcomes under profit-maximizing partitions

	No Communication		**With Communication**				
ϕ	(K_1^N, K_2^N)	Π_m^N	(x_l^*, x_h^*)	$(K_1^*(S_1), K_2^*(S_1))$	$(K_1^*(S_2), K_2^*(S_2))$	Π_m^*	$\frac{\Pi_m^* - \Pi_m^N}{\Pi_m^N}$ (%)
0.75	(245.1, 235.1)	147.3	(502.1, 502.7)	(217.4, 262.9)	(252.7, 227.6)	147.6	0.18
0.7	(251.1, 235.1)	188.1	(502.2, 503.4)	(191, 295.3)	(268, 218.3)	189.2	0.60
0.65	(256, 235.1)	235.5	(501.2, 502.7)	(123.7, 367.4)	(274.3, 216.8)	238.2	1.15
0.6	(260.7, 235.1)	291.1	(500.5, 502.9)	(16.9, 478.9)	(287.8, 208)	302.5	3.92
0.55	(265.3, 235.1)	357.3	(500, 505.6)	(16.7, 483.7)	(328.2, 172.2)	396.3	10.92
0.5	(269.9, 235.1)	437.2	(510, 500)	(382.4, 122.6)	(19.2, 485.8)	513.0	17.34
0.45	(274.5, 235.1)	535.4	(510, 504.4)	(337.4, 172.2)	(21.5, 488.1)	575.0	7.40
0.4	(279.1, 235.1)	658.7	(509.5, 507.1)	(307.2, 207)	(28.3, 485.9)	670.6	1.81
0.35	(283.8, 235.1)	817.9	(509.2, 508)	(298.7, 220.2)	(110.1, 408.8)	821.3	0.42
0.3	(288.6, 235.1)	1031	(509.2, 508.5)	(297.5, 226.2)	(184.8, 338.9)	1032.5	0.15
0.25	(294.6, 235.1)	1330.3	(509.2, 508.9)	(298.9, 230.8)	(244.2, 285.6)	1331.1	0.06

Note: $(r_1, c_1) = (1, 4.5)$ and $(r_2, c_2) = (5, 0)$, and w is varied such that $\phi = \frac{r_1}{w - c_1} \in \{0.75, 0.70, ..., 0.25\}$.

low-accuracy" messages) for $\phi > 0.5$, and $x_l^* > x_h^*$ ("low-average, low-accuracy" vs "high-average, high-accuracy" messages) for $\phi \leq 0.5$. The equilibrium capacity levels are tailored to the revealed accuracy information, i.e., more flexible capacity is reserved given a low-accuracy message. The difference $|x_h^* - x_l^*|$ and the value of communication attain the maximum at $\phi = 0.5$, and shrink as ϕ moves further away from 0.5 in either direction. Furthermore, we also compared the manufacturer's equilibrium profit under the profit-maximizing partition with that under the symmetric partition. We observed that the maximized profit is on average 1.37% higher than that under the symmetric partition. Therefore, symmetric partitions as discussed in §4 can be used as a good approximation if it is difficult for supply chain firms to pinpoint and conform to the specific profit-maximizing partition during communication.

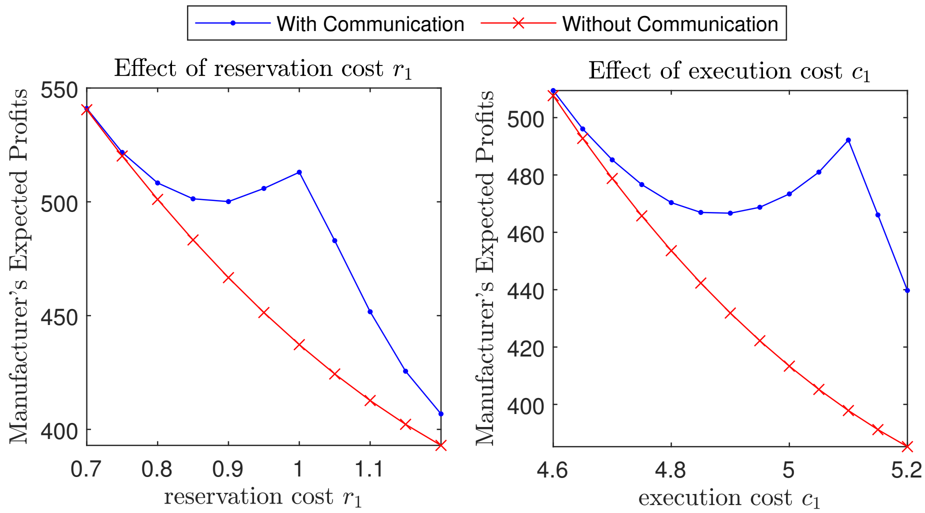

Figure F.1: The manufacturer's expected profit as a function of r_1 or c_1 under profit-maximizing partitions. In the left panel, r_1 is varied while $c_1 = 5$; in the right panel, c_1 is varied while $r_1 = 0.7$.

We also examined the effect of capacity costs on the manufacturer's expected profit under the profit-maximizing partition. We consider the same setting as described above but vary r_1 or c_1 with w being fixed at 6.5. As shown in Figure F.1, the manufacturer can be better off with an increase in r_1 or c_1, which confirms that our main findings in §4.3 are robust under profit-maximizing partitions.

G Effect of ρ_h on Profits

Recall that ρ_h represents the prior probability of the retailer having a high accuracy level. With a larger ρ_h, the retailer has better forecast capability ex ante, and so the supply chain faces less uncertainty in demand. In this appendix, we study how probability ρ_h influences the supply chain profits.

G.1 Retailer

We have proved that the manufacturer will choose the same total capacity level in any informative equilibrium as in the case of no communication. That is, the total capacity in equilibrium K_T^* is one that maximizes the manufacturer's expected profit under the prior distribution of (a, b). Because the retailer's payoff depends on the capacity portfolio only through K_T^*, whether or not informative communication occurs, the retailer will receive the same equilibrium payoff. Thus, analyzing the retailer's payoff under any informative equilibrium is equivalent to analyzing that without any communication. Nevertheless, as ρ_h increases, uncertainty in the prior demand distribution reduces such that the manufacturer may strategically adjust her total capacity level. As shown below, this can result in a nontrivial impact on the retailer's payoff.

To derive analytical results, we focus on the same setting as in Proposition 2(i), i.e., one with uniformly distributed a and ϵ with regularity conditions (1) and (2). The proof of Proposition G.1 can be found at the end of this appendix.

Proposition G.1. *Under the informative equilibrium characterized in Proposition 2, the following statements hold:*

(i) The total capacity in equilibrium is increasing in ρ_h if $\phi > \frac{1}{2}$, decreasing in ρ_h if $\phi < \frac{1}{2}$, and constant in ρ_h if $\phi = \frac{1}{2}$.

(ii) For $\phi \geq \frac{1}{2}$, the retailer's expected profit increases with ρ_h; however, for $\phi < \frac{1}{2}$, the retailer's expected profit can possibly decrease with ρ_h.

As shown in part (i) of Proposition G.1, the impact of ρ_h on the manufacturer's optimal total capacity K_T^* depends on the value of ϕ. Recall that the optimal total capacity is the solution to a newsvendor problem associated with an optimal service level of $1-\phi$. A larger ρ_h reduces the uncertainty in the prior demand distribution. When the optimal service level is exactly 1/2, the total capacity is equal to the average demand and constant in demand variability. For an optimal service level smaller than 1/2 (i.e., $\phi > 1/2$), the optimal capacity increases as demand becomes less variable, whereas for an optimal service level larger than 1/2 (i.e., $\phi < 1/2$), the optimal capacity decreases as demand becomes less variable.

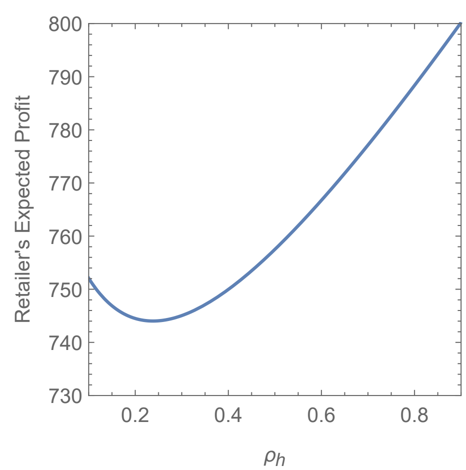

Figure G.1: The retailer's expected profit as a function of ρ_h; Parameter values: $p = 10, w = 8.5, r_1 = 1, c_1 = 4.5, b_h = 5, b_l = 1, a \sim U[500, 600], \epsilon \sim U[-500, 500]$

Part (ii) of Proposition G.1 characterizes the impact of ρ_h on the retailer's payoff. Interestingly, the retailer is not necessarily better off with a reduction in demand uncertainty. On the one hand, a larger ρ_h has direct value since the supply chain faces less uncertain demand. On the other hand, ρ_h influences the total capacity K_T^*, which indirectly affects the retailer's payoff. For the cases of $\phi \geq 1/2$, a larger ρ_h induces a higher total capacity level, and thus always benefit the retailer. However, for the case of $\phi < 1/2$, a larger ρ_h leads to a lower total capacity, leading to an indirect negative impact on the retailer's payoff. In the proof of Proposition G.1, we identify a range of parameters in which the negative indirect impact of a larger ρ_h dominates its direct value, thus making the retailer worse off. See a numerical example in Figure G.1, where an increase in ρ_h can reduce the retailer's expected profit unless ρ_h is large enough.

G.2 Manufacturer

We first write out the manufacturer's expected profit under uniformly distributed a and ϵ.

$$
\begin{aligned}
\Pi_m^* = \frac{(w-c_n)(\bar{a}+\underline{a})}{2} &- \Pr(S_1)\left(\sum_{i=1}^{n}(c_{i-1}-c_i)\sum_{t=h,l}\frac{\rho_t(x_t-\underline{a})}{\sum_{t'=h,l}\rho_t(x_t-\underline{a})}\int_{\underline{a}}^{x_t}\int_{-\sigma}^{\sigma}\frac{(a+\frac{\epsilon}{b_t}-\sum_{j=i}^{n}K_j^*(S_1))^+}{2\sigma(x_t-\underline{a})}d\epsilon da\right) \\
&- \Pr(S_2)\left(\sum_{i=1}^{n}(c_{i-1}-c_i)\sum_{t=h,l}\frac{\rho_t(\bar{a}-x_t)}{\sum_{t'=h,l}\rho_{t'}(\bar{a}-x_{t'})}\int_{x_t}^{\bar{a}}\int_{-\sigma}^{\sigma}\frac{(a+\frac{\epsilon}{b_t}-\sum_{j=i}^{n}K_j^*(S_2))^+}{2\sigma(\bar{a}-x_t)}d\epsilon da\right) \\
&- \Pr(S_1)\sum_{i=1}^{n}r_iK_i^*(S_1) - \Pr(S_2)\sum_{i=1}^{n}r_iK_i^*(S_2)
\end{aligned}
$$

where $\Pr(S_1) = \frac{\sum_{t=h,l}\rho_t x_t - \underline{a}}{\bar{a}-\underline{a}}$ and $\Pr(S_2) = \frac{\bar{a}-\sum_{t=h,l}\rho_t x_t}{\bar{a}-\underline{a}}$ denote the probabilities of $(a,b) \in S_1$ and $(a,b) \in S_2$, respectively. Substituting the expressions of $\Pr(S_1)$ and $\Pr(S_2)$, we can simplify the above into

$$
\Pi_m^* = \frac{(w-c_n)(\bar{a}-\underline{a})}{2} - \frac{\sum_{t=h,l}\rho_t(L_{t1}+L_{t2})}{2\sigma(\bar{a}-\underline{a})} - \sum_{i=1}^{n}r_i\left(\Pr(S_1)K_i^*(S_1)+\Pr(S_2)K_i^*(S_2)\right) \qquad \text{(G.1)}
$$

where, for $t = h, l$,

$$
L_{t1} = \sum_{i=1}^{n}(c_{i-1}-c_i)\int_{\underline{a}}^{x_t}\int_{-\sigma}^{\sigma}(a+\frac{\epsilon}{b_t}-\sum_{j=i}^{n}K_j^*(S_1))^+ d\epsilon da
$$

and

$$L_{t2} = \sum_{i=1}^{n}(c_{i-1} - c_i) \int_{x_t}^{\bar{a}} \int_{-\sigma}^{\sigma} (a + \frac{\epsilon}{b_t} - \sum_{j=i}^{n} K_j^*(S_2))^+ d\epsilon da.$$

In the expression of Π_m^*, x_l, x_h and the optimal capacity levels are all functions of ρ_h. In general, an increase in ρ_h has a direct value for the manufacturer as it reduces demand uncertainty, whereas ρ_h also influences the informative partition (i.e., x_l and x_h) and hence the specific capacity choices conditioning on S_1 and S_2. Unfortunately, a general analysis regarding the impact of ρ_h on Π_m^* is intractable. We therefore resort to a numerical study to explore the impact of ρ_h on the manufacturer's expected profit.

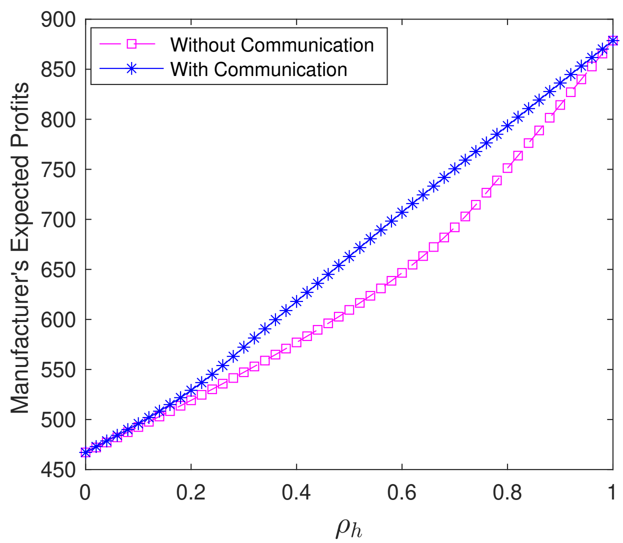

Figure G.2: The manufacturer's expected profits (with and without communication) as a function of ρ_h; Parameter values: $w = 6.5$, $(r_1, c_1) = (1, 4.5)$, $(r_2, c_2) = (5, 0)$, $a \sim U[500, 600]$, $b_l = 1$, $b_h = 20$, and ϵ follows a truncated normal distribution over $[-500, 500]$ with mean 0 and standard deviation 1000.

We consider two capacity sources and the following parameter values: $(r_1, c_1) = (1, 4.5)$, $(r_2, c_2) = (5, 0)$, $b_l = 1$, $b_h = 20$, and ϵ follows a truncated normal distribution over $[-500, 500]$ with mean 0 and standard deviation 1000. We vary the wholesale price w such that ϕ ranges from 0.1 to 0.8. Moreover, we consider uniformly distributed a but test two different supports: $a \sim U[500, 510]$ and $a \sim U[500, 600]$. For all the instances, we compute the manufacturer's expected profits with and without communication. We observe that the manufacturer's expected profit is increasing in ρ_h. Comparing the profits with and without communication, we also observe that the value of communication (i.e., the gap between the two profits) is maximized when ρ_h takes a moderate value, since the degree of information asymmetry in forecast accuracy shrinks as ρ_h becomes close to 0 or

1. A representative instance is presented in Figure G.2. Our numerical observations suggest that although ρ_h affects the value of communication in a non-monotonic manner, the direct benefit of increasing ρ_h always outweighs its indirect impact on the value of communication.

G.3 Proof of Proposition G.1

Proof of Proposition G.1. In the proof of Proposition 2, we have shown that the total capacity equals $K_T^* = \frac{\underline{a}+\bar{a}}{2} + \frac{\sigma(1-2\phi)}{\rho_h b_h + (1-\rho_h) b_l}$. Part (i) follows immediately by evaluating the derivative $\frac{\partial K_T^*}{\partial b_h} = \frac{(2\phi-1)(b_h-b_l)\sigma}{(\rho_h b_h + \rho_l b_l)^2}$.

For part (ii), first note that the retailer's expected profit can be written as

$$\Pi_r^* = (p-w) \sum_{t=h,l} \rho_t \mathbb{E}[\min(a + \epsilon/b_t + K_T^*)].$$

Then,

$$\begin{aligned} \frac{\partial \Pi_r}{\partial \rho_h} &= (p-w) \left(\mathbb{E}\left[\min(a + \frac{\epsilon}{b_h}, K_T^*)\right] - \mathbb{E}\left[\min(a + \frac{\epsilon}{b_l}, K_T^*)\right] + \frac{\partial K_T^*}{\partial \rho_h} \sum_{t=h,l} \rho_t \mathbb{E}[\bar{F}((K_T^* - a) b_t)] \right) \\ &= (p-w) \left(\mathbb{E}\left[\min(a + \frac{\epsilon}{b_h}, K_T^*)\right] - \mathbb{E}\left[\min(a + \frac{\epsilon}{b_l}, K_T^*)\right] + \phi \frac{\partial K_T^*}{\partial \rho_h} \right) \end{aligned}$$

where the last equality follows by the manufacturer's optimality condition $\phi = \mathbb{E}_{a,b}[\bar{F}((K_T^* - a)b]$.

The first two terms represent the direct benefit from the reduction in forecast error while the third term represents the indirect impact due to the manufacturer's strategic choice on K_T^*. The direct benefit (i.e., the first two terms) is nonnegative because of the properties of the second-order stochastic dominance. (That is, $\min(D, k)$ is concave in D and so $\mathbb{E}[\min(D_h, k)] \geq \mathbb{E}[\min(D_l, k)]$ if D_l is a mean-preserving spread of D_h. The indirect impact, i.e., the third term $\phi \frac{\partial K_T^*}{\partial \rho_h}$, is also nonnegative if $\phi \geq \frac{1}{2}$ by part (i). Thus, Π_r^* is increasing in ρ_h for $\phi \geq \frac{1}{2}$.

However, the indirect impact becomes negative for $\phi < \frac{1}{2}$. To prove the statement for $\phi < \frac{1}{2}$ in part (ii), it suffices to construct an example such that the indirect impact $\phi \frac{\partial K_T^*}{\partial \rho_h}$ dominates the direct benefit, so $\frac{\partial \Pi_r}{\partial \rho_h} < 0$. Notice that $\phi \frac{\partial K_T^*}{\partial \rho_h} = \frac{\phi(2\phi-1)(b_h-b_l)\sigma}{(\rho_h b_h + \rho_l b_l)^2}$ is minimized at $\phi = \frac{1}{4}$. Thus, we choose $\phi = \frac{1}{4}$, and by substituting the distribution functions of ϵ and a, $\frac{\partial \Pi_r^*}{\partial \rho_h}$ can be simplified to

$$\frac{\partial \Pi_r^*}{\partial \rho_h} = \frac{b_h - b_l}{48\sigma} \left[3\sigma^2 \left(\frac{4}{b_h b_l} - \frac{3}{(\rho_h b_h + \rho_l b_l)^2} \right) - (\bar{a} - \underline{a})^2 \right].$$

Note that for the case of $\frac{b_l}{b_h} < \frac{3}{4}$, $\frac{4}{b_h b_l} - \frac{3}{(\rho_h b_h + \rho_l b_l)^2} < 0$ as $\rho_h \to 0$ (and $\rho_l \to 1$). Therefore, for small enough ρ_h, Π_r^* is decreasing in ρ_h in the case of $\phi = \frac{1}{4}$ and $\frac{b_l}{b_h} < \frac{3}{4}$.

□

H Discussion on the Contract Forms

An important feature of our model is that *the retailer's payoff depends on the manufacturer's capacity portfolio only through the total capacity level under the given supply chain contract.* As such, the supply chain payment needs to be based on the total units delivered and the demand, irrespective of how much of each capacity source is used (i.e., the portfolio breakdown). This is true under wholesale price contracts as we assumed in the paper. Nevertheless, there are other forms of contracts that retain this feature.

- **Revenue Sharing.** If the retailer agrees to share an α portion of its revenue to the manufacturer, then in our setting where the retailer's actual order is placed after demand realization, it is equivalent to having the wholesale price as $w = \alpha p$.

- **Under-delivery Penalty.** In addition to the wholesale price w for each delivered unit, the manufacturer may pay the retailer an under-delivery penalty u for each unit of unsatisfied demand. In this case, for any reserved capacity levels (K_1, K_2), the manufacturer's second-stage profit for capacity execution becomes

$$\begin{aligned} & w \min(x_1 + x_2, D) - u(D - x_1 - x_2)^+ - \sum_{i=1}^{2} c_i x_i \\ =& (w+u) \min(x_1 + x_2, D) - uD - \sum_{i=1}^{2} c_i x_i \end{aligned}$$

  where the equality follows by the relation $(x-y)^+ = x - \min(x, y)$. Hence, the manufacturer solves the same problem except that it treats $w + u$ as the original wholesale price w in our model. On the other hand, the retailer's expected profit is given by

$$\Pi_r(K_1, K_2) = (p - w)\mathbb{E}[\min(D, K_1 + K_2)] + u\mathbb{E}[(D - K_1 - K_2)^+],$$

  thus depending on the capacity portfolio only through the total capacity level.

- **Excess Capacity Subsidy.** In some industries, the retailer may compensate the manufacturer for the unused capacity. Let s denote the subsidy that the retailer pays to the manufacturer for each unit of unused capacity, regardless of which capacity source. Then,

$$\begin{aligned}&w\min(x_1+x_2,D)+s(x_1+x_2-D)^+-\sum_{i=1}^{2}c_ix_i\\=&(w-s)\min(x_1+x_2,D)-\sum_{i=1}^{2}(c_i-s)x_i\end{aligned}$$

As long as $s<c_i<w$ for all i, the manufacturer's problem will have the same structure as in our model while the original wholesale price w is replaced by $w-s$ and the original execution cost c_i is replaced by c_i-s. Then, the retailer's expected profit is given by

$$\Pi_r(K_1,K_2)=(p-w)\mathbb{E}[\min(D,K_1+K_2)]-s\mathbb{E}[(K_1+K_2-D)^+],$$

which depends only on K_1+K_2.

Therefore, under the aforementioned contract forms, our results continue to hold because the retailer cares only about the manufacturer's total capacity level. Having said that, there can be other forms of contracts under which the capacity portfolio breakdown matters to the retailer. For instance, supply chain firms may agree to share the profit or cost, instead of sharing the revenue. The retailer may subsidize the unused capacity based on specific capacity source. In these cases, the retailer's profit will depend on the specific values of K_1 and K_2 such that our results cannot carry over.

I Noncontiguous Partitions with Nonconvex Sets of a

In the main body of the paper, we focus on the partitions such that within S_1 and S_2, the set of a is convex on each accuracy level. Theoretically, there can be noncontiguous partitions that contain nonconvex sets, although such partitions seem to have little managerial interpretation. For any $0<x<\frac{\bar{a}-\underline{a}}{2}$, we consider a *noncontiguous* partition $S_1(x)=\{(a,b)\in S|a\leq\underline{a}+x \text{ or } a>\bar{a}-x\}$ and $S_2(x)=\{(a,b)\in S|\underline{a}+x<a\leq\bar{a}-x\}$. As illustrated in Figure I.1, S_1 contains a nonconvex set of a on each accuracy level.

Consider the special case of uniformly distributed demand average and noise with equally likely accuracy levels: $a\sim U[\underline{a},\bar{a}]$, $\epsilon\sim U[-\sigma,\sigma]$ and $\rho_h=\rho_l=0.5$. We show that the $S_i(x)$'s constructed

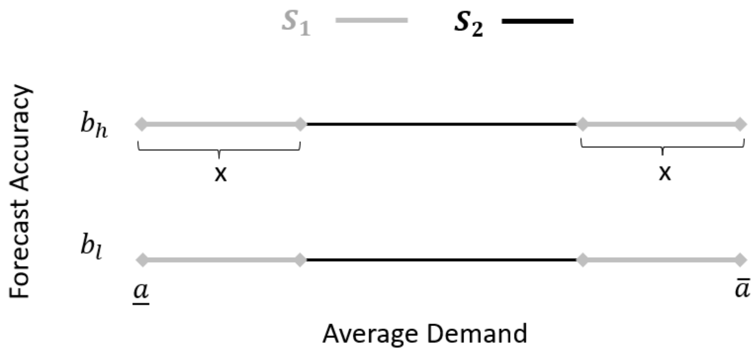

Figure I.1: Illustration of noncontiguous partitions

above constitute an informative partition for $\phi = 0.5$. In this case, the optimal total capacity without communication is given by $K_T^N = \frac{\underline{a}+\bar{a}}{2}$. It then suffices to show that K_T^N is also optimal conditioning on $S_1(x)$ and $S_2(x)$, respectively.

Conditional on $S_1(x)$, the optimality condition for the total capacity is written as

$$\frac{1}{4x}\sum_{t=l,h}\left(\int_{\underline{a}}^{\underline{a}+x}\bar{F}((K_T-a)b_t)da+\int_{\bar{a}-x}^{\bar{a}}\bar{F}((K_T-a)b_t)da\right)=\frac{1}{2}. \tag{I.1}$$

Conditional on $S_2(x)$, the optimality condition for the total capacity is written as

$$\frac{1}{2(\bar{a}-\underline{a}-2x)}\sum_{t=l,h}\int_{\underline{a}+x}^{\bar{a}-x}\bar{F}((K_T-a)b_t)da=\frac{1}{2}. \tag{I.2}$$

One can show that $K_T^* = \frac{\underline{a}+\bar{a}}{2}$ solves both (I.1) and (I.2), provided that $\frac{\underline{a}+\bar{a}}{2} \in (a-\frac{\sigma}{b}, a+\frac{\sigma}{b})$ for all $(a,b) \in S$. Hence, $S_1(x)$ and $S_2(x)$ constructed above support an informative equilibrium.