NORTHWESTERN UNIVERSITY

# Supply Chain Operations with Partial Demand Information and Customer Behavior

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Industrial Engineering and Management Sciences

By

Pol Boada-Collado

EVANSTON, ILLINOIS

June 2021

# ABSTRACT

Supply Chain Operations with Partial Demand Information and Customer Behavior

Pol Boada-Collado

This dissertation is motivated by the decision process of the supply chain team of a major furniture retailer that delivers its products at the customer's home. In retail supply chains for companies offering home delivery services, demand surges are observed at the store level, which translate to a high volume of orders that need to be delivered. Ordered products are distributed through middle mile lanes that connect hubs to spokes, from where the last-mile delivery routes depart. This dissertation studies novel strategies to reduce distribution costs by adapting middle and last mile operations to changes in the market place such as demand surges and consumers' buying trends.

Chapters 2 and 3 of this dissertation focus on the middle mile transportation procurement of hub-to-spoke lanes when signing transportation capacity contracts with short temporal commitment. With new technological innovations, short temporal commitment contracts are found in dynamic environments like distribution, processing, and manufacturing; a trend likely to grow in the future. In contrast to classic procurement, where

commitments are long, short temporal commitments lead to new dynamics in which demand visibility at the spoke level (resulting from in-store orders and deliverable windows) fundamentally changes contracting policies. By studying first a single dedicated hub-to-spoke lane, and then extending the results to a distribution network with multiple lanes, we show that dynamically incorporating demand visibility in procurement decisions allows companies to counteract demand uncertainty more effectively, better respond to seasonal demand surges, and use transportation capacity more efficiently.

The settings studied in these chapters assume that delivery dates to customer homes are fixed. The last part of this dissertation, Chapter 4, introduces flexibility in delivery dates and focuses on scheduling policies for last mile distribution, i.e.; from spokes to the customer's home, to mitigate demand surges at the spokes level. When planning middle and last mile transportation capacity, understanding customers' sensibility to delivery lead time allows retailers to balance customer satisfaction and shipping costs. In attended home delivery, where customers need to be present at the time of delivery, as in the case of the furniture company, we study the importance of speed of delivery relative to other priorities when customers choose their preferred home delivery date. We find that speed of delivery is not only the factor that matters to customers, and it is of limited importance relative to customer availability and day of the week preferences. For the retailers, this is an opportunity to slightly extend delivery windows and more efficiently utilize middle and last mile distribution capacity.

Taken together, the chapters in this dissertation show how companies can adapt operations to reduce distribution costs while meeting the expectations of their customers.

## Acknowledgements

I would like to express my most sincere appreciation to my advisors Karen Smilowitz and Sunil Chopra. Thank you, Karen, for welcoming me as your advisee, and supporting me throughout the Ph.D., not only academically, but far beyond. Your guidance has taught me to chase relevant and impactful challenges, and to overcome the complexities and difficulties of my projects. Also, I am thankful for your cheerfulness, and for instilling in me values to make the world a better place. Thank you, Sunil, for conveying your passion for research, and for teaching me to identify the key questions and the main insights, and to keep always the big picture in mind. Also, thank you for sharing with me your wisdom and advice for effectively communicating and disseminating. I would also like to extend my deepest gratitude to both of you for your optimism towards me, your patience, your understanding, and for unconditionally supporting me during my job market. I am the luckiest student for having such two wonderful advisers.

I would like to extend my sincere thanks to my dissertation committee, Maria Ibanez, Simge Küçükyavuz, and Jan A. Van Mieghem, for their time, effort, valuable comments, questions, and suggestions, which have greatly improved the quality of my dissertation.

I am deeply indebted to Victor Martinez de Albeniz, who believed in me when I barely had any research experience. Victor taught me, encouraged me, and pushed me to shoot for the stars; he made possible this adventure. I also wish to thank Maria Victoria Sanchez-Vives for welcoming me into her laboratory, for hearing my research ideas, for believing in me, and for encouraging me to pursue the Ph.D. I am also grateful to the Department of Industrial Engineering and Management Sciences at Northwestern University, whose faculty and staff believed in me and offered me a place to start my journey. Also, many thanks for the invaluable help I received from the department during the Ph.D. I gratefully acknowledge the financial support during my last year from the Dissertation Terminal Year Fellowship of the Northwestern University, and the NUTC Dissertation Year Fellowship of the Northwestern University Transportation Center.

I would also like to thank the current and former Karen's team members, Mariana Escallón Barrios, Min Fei, Reut Noham, Aysu Özel, Eugene Wickett, and Liwei Zeng, for hearing my presentations uncountable times, for their feedback, and for making it the best team to work with. Many thanks to a great office mate and friend, Daniel Duque, for our great conversations, whiteboard sessions, and laughter inside and outside the office. Also, thanks to all my friends from the Northwestern University, Ghani Ebrahimi, Linda Pei, Suraj Yerramilli, and many others, for making the Ph.D. a wonderful journey.

I would like to extend my deepest gratitude to my family, who have always supported my education. Since I was a child, my parents and my brother have taught me to be curious, and have always valued my creativity and enthusiasm, which are qualities that have been crucial throughout these years. I would also like to extend my gratitude to my wife's family for always encouraging us to reach our goals. To our families, thank you so much for always believing in us, and for your unconditional support.[1]

[1] *In Catalan:* Voldria expressar el meu agraïment sincer a la meva família, que sempre m'ha recolzat en els meus estudis. Des de que era petit, els meus pares i el meu germà m'han ensenyat a ser curiós, i han valorat la meva creativitat, il.lusió i entusiasme, que són qualitats que m'han portat fins aquí. També vull agrair a la família de l'Anna, que sempre ens ha animat a aconseguir allò que ens hem proposat. A les nostres famílies, moltes gràcies pel vostre suport incondicional i per sempre creure en nosaltres.

Finally, yet importantly, I would like to thank my wife, Anna, with all my heart. Anna, none of this could have been possible without your help, your revisions, your valuable advice, and your patience. You are my model of strength, perseverance, thoroughness, and accuracy, some of the qualities that make you greatly succeed in your Ph.D. Thank you for being so understanding, for believing in me, for cheering me when I needed it the most, and for all the laughter. Thank you for your company during the good times and during not-so-good times. Thank you for making me want you to feel proud of me. Thank you for making possible our adventure, which has only just begun.

*To my wife, Anna, and my nieces and nephew, Abril, Carla, Pere and Pia.*

# Table of Contents

# List of Tables

# List of Figures

CHAPTER 1

# Introduction

Motivated by a collaboration with the supply chain team of a major furniture company, this dissertation studies novel strategies to improve middle and last mile operations of retail companies that deliver their products to the customer's home. Home delivery has grown rapidly with the emergence of e-commerce [**94**], and increasing customers' expectations and the competing pressure of big retailer companies such as Amazon, Alibaba, or Walmart has pushed retailers to assume the entire cost of transportation [**98, 70**]. Hence, reducing the cost of distribution across the supply chain has become crucial for all retailer companies aiming to stay competitive.

With the growth of flexible technologies in supply chains, many traditional static supply models have become obsolete. As industry moves to more dynamic environments and market places become more competitive, supply chains are evolving accordingly by adapting to market changes, such as demand surges or consumers' buying trends. The need for more agile supply networks is even more evident during pandemic times. This dissertation studies how middle mile and last mile distribution, two of the most expensive and complex operations of the supply chain, can become more efficient and nimble by adapting transportation procurement policies to demand surges and customers' expectations.

The furniture company that motivates this dissertation, which operates with more than 550 locations across the United States, delivers its made-to-order products to the customer's home. As sketched in Figure 1.1, this operation, which represents a main logistic challenge for the supply chain team, requires the distribution of products from hubs to distribution centers, from where the last mile delivery routes depart. To distribute the products from the hubs to the distribution centers, the company outsources its middle mile transportation capacity by signing transportation procurement contracts. Chapters 2 and 3 focus on the middle mile operations of the furniture company, and show how transportation procurement policies for middle mile lanes can adapt to demand fluctuations observed at the distribution centers.

Once products arrive at the distribution centers, delivery teams that work for the company carry out the lat mile distribution to the customer's home. In contrast to the middle mile transportation capacity, the last mile transportation capacity is owned by the company. Chapter 4 focuses on the last mile distribution, and, using transactional data from the furniture company, shows that customers' priorities when choosing home delivery dates can be leveraged to use more efficiently the last mile delivery capacity.

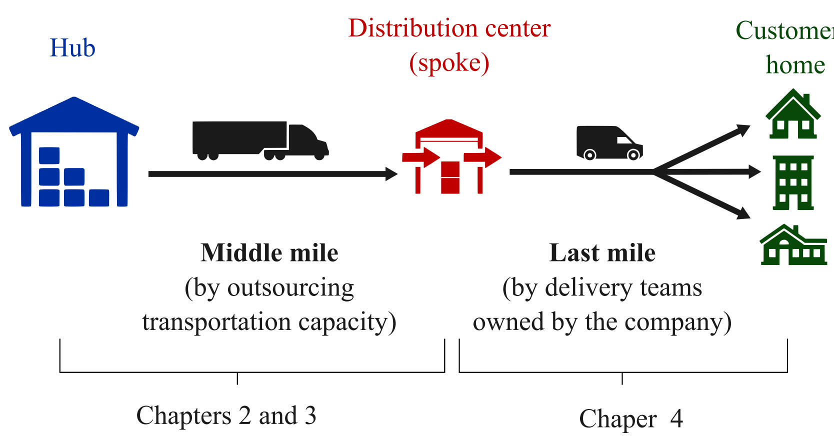


Figure 1.1. Structure of the dissertation.

## 1.1. Middle Mile Operations: Demand Visibility in Procurement with Temporal Commitment

Chapters 2 and 3 focus on adapting middle mile operations to demand visibility, a scenario motivated by the decision process faced by the supply chain of the furniture company. Due to delivery time windows (typically one or two weeks), line-haul transportation needs from hub to spokes are visible one week in advance, which can be supplied by signing transportation procurement contracts at a chosen capacity with relatively short commitment (one month). In these two chapters, we study how information from one week of demand visibility can be leveraged when signing procurement contracts with relatively short temporal commitment.

In traditional procurement, commitments are typically long, spanning several months or years, because deployment of capacity is expensive and cannot be changed regularly. In contrast, in transportation procurement, capacity is more easily modified by changing the number of trucks assigned to a lane, allowing for shorter commitments. Demand visibility for a few weeks becomes valuable with short commitment contracts. With technological innovations, our focus on short term commitments extends to dynamic environments in logistics and manufacturing with easily re-deployable capacity [**45**]. For instance, Flexe is a company that provides on-demand warehousing and allows commitments as short as a few weeks. Similarly, in manufacturing, 3D printing provides firms with a production capacity that is easy to redeploy.

Chapter 2, which is based on [**19**], provides foundational insights into the use of demand visibility for signing procurement contracts with temporal commitment. In a single hub-to-spoke lane, we model the decision process of the furniture company as a Markov

Decision Process and show demand visibility fundamentally changes the contracting policies, increasing the decision-makers' willingness to commit to transportation procurement contracts. When committing to a contract, the optimal capacity balances the trade-off between adapting to the observed demand and remaining cost-efficient for future periods, for which demand is not visible. Using data containing sales transactions of the furniture company, we show that, with seasonal demand, decision-makers leverage demand visibility by coordinating the contracting epochs with expected demand shocks (such as Black Friday or Labor Day holidays sales).

Chapter 3, which is based on [**20**], extends the analysis of the first chapter to a network with multiple distribution centers (or nodes). In a network, flexible transportation capacity that can supply multiple nodes offers an additional hedging mechanism against demand uncertainty (in addition to demand visibility) when signing contracts with temporal commitment. In contrast to node-dedicated transportation contracts, flexible contracts allow companies to pool transportation capacity between nearby distribution centers. With demand visibility and commitment, flexible capacity is most effective when observed demand of the nodes have a large difference with high observed demand at one node and low observed demand at the other node. Adapting to observed demand can lead to an increase of the average commitment to flexible capacity due to demand visibility, even in situations in which flexible capacity would be disregarded without visibility. Previous studies show that, without temporal commitment, the two mechanisms are substitutes. In this chapter we investigate and show how commitment imposed across periods leads to new dynamics in which the two hedging mechanisms act as both substitutes and complements.

## 1.2. Last Mile Operations: Adapting to customers' choice for attended home delivery

The first two projects assume that delivery lead time (which translates to demand visibility) is fixed exogenously. The next chapter explores the company's scheduling process of last mile delivery, and studies the impact of increasing or decreasing delivery lead time on customer satisfaction.

Once a sale is closed at the retail outlets, the company commits to deliver the purchased products at the customer's home on a date chosen by the customer. Once an order is entered into the system, products are prepared and distributed from a hub to the distribution centers, from where the last mile delivery routes depart. Chapter 4, which is based on [**18**], focuses on adapting last mile delivery scheduling policies to customer behavior patterns observed by the supply chain team of the furniture company. A large body of the literature, and a trend followed by practitioners, has focused on shortening delivery lead times as much as possible to increase customer service level [**70**]. Such an effort usually comes at high costs for companies and the environment [**42, 104**]. In attended home delivery, where customers must be present at the time of delivery, retailers need to quantify customers' sensibility to delivery lead time to balance customer satisfaction and shipping costs.

Using transactional data from the furniture company that motivates this dissertation, we model how customers choose attended home delivery dates from among the set of alternative dates offered by the company. The model allows us to estimate the customers' priorities that drive their choice and the relative impact of speed of delivery. The empirical results show that customers do not usually schedule attended home delivery for the days immediately after the purchase date. Speed of delivery is of limited importance relative to other priorities, such as day of week preferences or availability, when customers choose their preferred delivery date. For companies committed to deliver their products, this is an opportunity to slightly extend delivery windows and more efficiently utilize transportation capacity. Also, the empirical results show that customers have strong day of the week preferences in attended home delivery, which can potentially be leveraged to decrease delivery lead times while reducing transportation capacity.

Our results show that companies committed to deliver their products at the customer's home can increase utilization of costly delivery capacity while adapting last mile operations to meet customer's expectations.

The last chapter, Chapter 5, presents open questions and future research directions on the exploration of novel strategies to reduce distribution costs in middle and last mile operations.

CHAPTER 2

# Partial Demand Information and Commitment in Dynamic Transportation Procurement

*We investigate the value of short-term demand visibility in the decision process of selecting the optimal transportation procurement in single-leg distribution lanes. Our work considers procurement contracts with temporal commitment horizon that is much shorter than the planning horizon, a scenario motivated by the supply chain team of the furniture retailer. We show that having this demand information essentially changes contracting policies by adding more dynamism and, with seasonal demand, information is leveraged by properly coordinating with expected demand shocks (e.g. Black Friday) using tailored strategies.*

## 2.1. Introduction

Selecting committed contracts or more expensive spot market options to meet transportation needs is a critical trade-off often faced in industry. Given the significance and prevalence of this decision, it has been well studied by the operations research community (see the review of [**115**] on multimodal transportation and the work of [**79**] and [**102**] on modal rate estimation). In this chapter, we study how the availability of partial future demand information impacts this decision, motivated by the transportation needs of a major furniture manufacturer.

As sales are registered at retail outlets and future delivery dates are finalized, the transportation group of the furniture manufacturer typically has visibility into scheduled delivery needs over the next week. Accordingly, the demand for transportation along lanes is known one week in advance. On any lane, the company has the option to fulfill transportation needs for the coming week using the spot market through services such as XPO, UPS, or USPS. The company also has the option of signing a longer term contract (typically three to four weeks) for a committed capacity. The contract has a lower unit price than the spot market. With one week of demand visibility, the company must decide whether or not to sign a contract, and, if so, at what capacity. The company asked the following question: How can the limited visibility regarding future demand be used when making transportation decisions in the presence of commitment?

This decision problem has two key features that, when combined, distinguish this study from the literature. The first feature is the availability of *partial demand information*, providing demand information for the early part of the planning horizon. In other words, the partial demand information setting implies demand visibility of a few periods in the future (one week for the case of the furniture company). A large body of literature has focused on the *full information* case where demand for the entire planning horizon is known (see for instance [**25**] and [**36**] for two reviews on lot sizing models) or the *null information* case where decisions are made before observing future demand in the planning horizon but a demand distribution is known (e.g. [**12**]). The second feature is that the *commitment* contract length in our setting has a horizon that is limited and shorter than the planning horizon. Existing literature has largely focused on commitment contracts where the commitment horizon equals the entire planning horizon, which is well characterized in [**8**] with full information model and in [**93**] with null information.

We show that these two features, when studied together, result in an optimal behavior that differs from the full and null information cases. When full information is available about demand in the planning horizon, mixed integer programming models determine the optimal use of commitment contracts and the spot market over the entire planning horizon [**8**]. In contrast, when no future demand information is known, supply contracts can be determined based on the expected performance over the planning horizon [**93**]. In both cases, contract selection can be fully determined at the beginning of the planning horizon (in the case of full information this includes periods of no commitment and changes in capacity levels with new contracts), even though the quantities shipped are dynamic based on actual demand. The availability of partial demand information and a limited commitment horizon results in commitment decisions that change over time as new partial demand information becomes available, giving additional dynamism to the contracting policies.

This chapter provides foundational insights in the analysis of the role of partial information in the presence of short commitment. We begin with a scenario with a single lane that faces stochastic demand with a known distribution when dynamically selecting the best transportation modes. The decision maker must select the best choice among two options: signing a capacity contract that will remain in effect over the commitment horizon or going to the spot market for the entire demand for the coming period, which allows the decision to be revisited at the end of the next period. Once the decision is made, costs are incurred and the decision process resets to its initial state (either in the next period if no contract is signed or at the end of the commitment horizon if the contract is signed). Under stationary demand, the null and full information settings have simple policies that can be determined at the beginning of the horizon. When partial information is available and commitment length is short relative to the planning horizon, more complex policies emerge. If the contract has only one choice of capacity, the availability of partial demand information can lead to the signing of contracts that would not be signed under the null information scenario. When a decision maker can select among a wide range of capacities, the availability of partial information impacts both the decision of whether to sign the contract and the capacity level at which the contract is signed. The optimal capacity balances the trade-off between adapting to the observed demand and remaining cost-efficient for future periods for which demand information is not available. We extend our model to the case of seasonal demand along a lane, which is the scenario faced by the furniture company. We find that in the presence of seasonality, partial demand visibility can be leveraged by properly coordinating the contracting policy. Depending on the nature of the seasonality, the coordination should follow one of two strategies that we identify as *Wait-and-See* or *Now-or-Never*.

The chapter is structured as follows: Section 2.2 reviews the related literature. We describe the problem and the main results in Section 2.3. We summarize the conclusions in Section 2.4.

## 2.2. Literature review

There is an extensive body of literature that studies transportation mode choice in freight distribution (see [**138**] and [**115**] for two reviews of multi-modal freight transportation). There also exists a significant stream of literature that considers the valuation of transportation contracts, bidding mechanisms and financial aspects of the spot market (see for example [**126**], [**56**], [**4**] and [**113**]). In contrast to the cited articles, our work focuses on the cost of commitment and the value of information when the decision maker is faced with selecting the most cost-efficient contracts to meet their transportation needs. [**97**] qualitatively reviews the role of flexibility in logistic multi-modal transportation and provides a taxonomy that identifies different sources of flexibility. Our analysis adds a new dimension by considering how policies change when decision makers have limited demand visibility in the presence of commitment.

If one views the transportation contracts studied in this chapter more broadly as supply contracts, our work is related to the supply procurement contract and inventory management literature. We refer to [**6**] for a detailed description of the general framework of supply contracts. Close to our analysis, [**12**] develops a dynamic model that allows capacity commitments to be updated after demand is realized. Their problem setting differs from ours in two main dimensions: (1) decisions are made without observing demand of future periods and (2) the length of the contract is equal to the planning horizon but capacity levels can be updated. [**127**] studies the cost of commitment as a penalty incurred every time the decision maker uses less than the committed quantity. In our case, the spot market provides flexibility at higher prices, that can be used if demand exceeds the committed capacity or if the decision maker chooses to forgo a commitment contract. [**66**] provides a comprehensive review of supply contracting models that incorporate the spot market and analyze a model of supply procurement contracts that allows abandonment, which is conceptually close to having a planning horizon longer than contract duration, yet their setting lacks the periodic nature of repeated contract decisions.

Information in supply chain contracting has been modelled in various contexts (see [**111**] for a review). In this chapter, we distinguish three levels of information: Full information, null information and partial information. Under the full information scenario, demand is known for all periods in the planning horizon. Capacity decisions are studied using deterministic models where commitments are signed at the beginning of the planning horizon. Full information over the planning horizon is studied in [**8**]. The authors characterize the optimal solutions for capacity acquisition and subcontracting problems for the case where capacity is continuous with general concave cost functions. The authors formulate the problems with linear programs and develop algorithms that find the optimal in polynomial time. Other formulations related to capacity planning in manufacturing are reviewed in [**92**]. Some articles study complexity and polyhedral properties of lot sizing problems under full information as in [**9**]. Another example with full information is [**110**], where the authors analyze subcontracting policies when capacities are fixed and demand is known at each period. The work focuses on the complexity of the algorithms and provides solutions using dynamic programming. Our approach is different in considering a dynamic decision process where, at each decision epoch, demand is visible for a short horizon (information horizon). Beyond that horizon, only the probability distribution of the demand is known. In this scenario, commitments are dynamically adjusted depending on what is observed for the periods where demand is visible. If the information horizon is very long (approaching the length of the planning horizon), the decision maker may sign all commitments at the beginning of the planning horizon and our model would be equivalent to a special case of the model studied in [**8**], where capacities and commitments are deterministically fixed. The focus in this chapter, however, is on instances in which the information horizon is not long (a week or two), a situation that often arises in practice, distinguishing our work from full information variants. In the next paragraph, we discuss how the combined presence of short commitments and partial demand visibility distinguish our setting from related stochastic variants (null and partial information).

There is significant work on the null information case where no demand information (other than the distribution) is known over the planning horizon. In [**11**] an initial commitment, which fixes a total minimum inventory purchased periodically during the planning horizon, must be defined before observing the demand. [**119**] provides another example where decisions are made before observing the demand. In that case, the buyer commits to buy a total amount along the planning horizon. The authors in [**93**] propose an approach that mixes different contracts when demand and spot market prices are stochastic. In their setting, the contracts and commitment levels are chosen when demand is not known but order quantities are decided periodically after demand and spot market price realizations. Once the capacity decision is made, this cannot be changed or traded in later periods. In contrast, in our model contracts and commitment levels are chosen with partial demand visibility and can be altered at the end of each commitment horizon. [**73**] also finds solutions with random demand and spot market prices. In contrast to our model, demand is not realized at the time of ordering, only spot market price, and capacity reservation is decided at the beginning of the time horizon.

The models cited above assume very long commitments where capacity and costs of procurement contracts are negotiated long before the beginning of the season. This framework is suitable for industries where capacity installation is expensive and cannot be easily adjusted. In contrast, in the context of transportation procurement, capacity can be increased or decreased by changing the assignment of trucks dedicated to a lane. While this cannot be performed daily, it is possible to alter the capacities selected over a slightly longer commitment horizon (four weeks in the case of the furniture company). With relatively short commitments, we show that demand visibility is valuable. In our setting, the capacity of the contract is chosen each time a new commitment is made and, with demand visibility, the committed capacity may differ every time the decision maker signs a new commitment. In contrast to the models cited above, contracts are revisited periodically during the planning horizon. Note that when commitments are very long, demand visibility is not as valuable and our problem can be formulated as a single capacity decision that is made at the beginning of the commitment horizon. In this case, our model becomes a special case of the models cited above where contract features are set for the entire planning horizon (e.g. vector of capacities defined by [**93**]).

Partial demand information is modelled differently depending on the context, see [**132**] for a review of the value of information in supply chain decisions. They classify the literature that study the value of information in supply chains from 2006 to 2017. In the context of logistics and transportation, [**120**], [**145**] and [**146**] analyze the benefit of advanced load information for truckload carriers. The authors model a dynamic pick up and delivery truckload model in which orders are sequentially revealed. The orders are known some periods in advance, the authors call this visibility the "knowledge window".

They show that increased demand visibility can significatively increase profits. Our work is different in considering the value of information under committing decisions. [**147**] uses a stylized single-period single-lane model to study the value of information of container release times on determining volume and departure time in maritime container transport. While addressing similar questions, our model provides a different approach (dynamic decision process) in a different context (contracting transportation procurement under commitment). Linking the information of the demand to the temporal dimension allows to study its value in a (quasi-)continuous fashion.

In a more general context, [**78**] and [**77**] review inventory management models across different information scenarios. In their work, capacity commitment is not included in the decision and the spot market does not play the role of adding flexility. In a similar context, [**49**] and [**50**] study the impact of asymmetric demand information on the interaction between a seller and buyer. [**106**] models partial demand information as receiving a fraction of early purchases that will be supplied after the capacity and pricing decisions are made. Our work is different in considering information as demand realizations that are revealed some periods before the decision is to be made in a multiperiod setting. [**135**] develops a dynamic inventory management model with advance demand information and flexible fulfillment: Demand realizations are known some periods in advance and known demand realization can be fulfilled at any time. [**15**] studies the impact of advance demand information in a multiproduct system, where information can be either about the aggregate demand or about the market share of the individual products. In contrast, our work considers a multiperiod setting where demand realizations of early periods are known in advance. Finally, there is growing literature on the use of advance demand information (perfect or imperfect) for updating forecasts: Recent work in this area include the work in [**5**], where the authors study when advance demand information is most valuable in large scale assembly lines. [**103**] shows that advance demand information, when released sequentially, can improve gross profit more than 5%, and [**123**] studies the value of imperfect information with a dynamic inventory model. Our approach adds to this work by showing the value of information in contract decisions for shorter commitment durations.

### 2.3. Transportation Procurement: Partial Information Model

We consider the following setting faced by a major furniture manufacturer with a workforce of over 5000 employees and over 600 locations across the United States. The transportation needs of the company require regular movement of products from hubs to spokes. The periodic demand of products to be moved for each hub-spoke pair (a lane) is random and subject to change from one week to another, although historic demand is available to characterize the demand distribution for each lane. The system used for processing orders allows the company to schedule final delivery four or five days, or even a week, after the order is entered into the system, providing supply chain managers with demand visibility over the coming week. This time period is referred as the *information horizon*. Managers can potentially take advantage of this demand visibility by going to the spot market and obtaining the exact capacity needed to meet the upcoming demand. However, there are cost and complexity advantages to signing a longer term transportation contract. When signing such a contract, managers commit to a capacity level for a period of four weeks. This period is referred as the *commitment horizon*. Hence, the commitment horizon requires that each time a contract is signed, managers must account for the expected cost over the part of the commitment horizon when demand is not visible. When signing a contract, transportation needs in each period up to the committed capacity are served with the less expensive committed capacity, and transportation needs in excess of the committed capacity are obtained from the spot market at a higher cost.

We begin with a single lane that faces a random demand that is stationary, independent and identically distributed (iid), and extend the results to seasonal demand.

The decision process, sketched in Figure 2.1-(a) for an information horizon of one period, is as follows: at the end of each period $t$, decision makers encounter one of two situations. (1) There is a previously signed contract in place whose commitment horizon includes the upcoming period or (2) there is no active contract for the upcoming period and, hence, transportation procurement must be determined. Decision makers have an information horizon of $\gamma$ periods and available long term contracts have a commitment horizon of $\tau$ periods. Thus, the exact demand of periods $t+1, t+2, t+3, \ldots, t+\gamma$ is observed (denoted $d_t$ in Figure 2.1-(a)). We assume that the demand for consecutive periods is iid with a known stationary distribution (denoted $D_t^w$ in Figure 2.1-(a)). The decision maker chooses one of two options:

- **Contract Mode**: Sign a commitment contract with a chosen per-period capacity $k$ at cost $c$ per unit (total contract cost of $ck$ per period). The contract has a commitment horizon of $\tau$ periods. Hence, for each of the next consecutive periods $t+1, t+2, \ldots, t+\tau$, a fixed cost of $ck$ is incurred even when demand is less than $k$. If the demand in any period exceeds the contracted capacity, the decision maker uses the spot market to supply the excess demand at cost of $c'$ per unit of demand. At the end of period $t+\tau$ the decision process is reset. When this option is chosen, we say that the system is using a *contract mode* (denoted $m = 1$) that remains active from periods $t + 1$ to $t + \tau$.

- **Spot Market Mode**: Use the spot market at cost $c'$ per unit to meet the demand for period $t + 1$. At end of period $t + 1$, the decision process resets. Thus, the spot market option allows the decision maker to delay the decision of signing the contract by one period. When this option is chosen we say that the system is using the *spot market mode* (we denote $m = 0$) that remains active only for period $t + 1$.

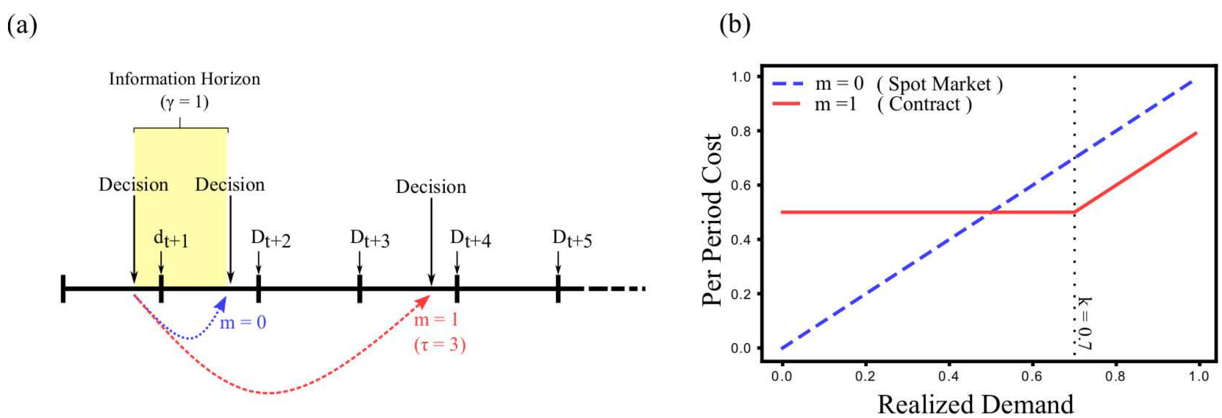


Figure 2.1. (a): Decision sequence for periods $t, \ldots, t + 5$. (b): Cost per period of each mode with $k = 0.7$, $c = 2/3$ and $c' = 1$

The cost in a period with demand $d$ is referred as $C^0(d)$ if the spot market mode ($m = 0$) is selected and $C^1(k, d)$ if the commitment mode ($m = 1$) at capacity $k$ is selected. We thus obtain the following:

$$
C^0(d) = \quad c'd
$$
$$
\tag{2.1} C^1(k,d) = \quad ck + c'[d-k]^+,
$$

where the operator $[x]^+ = \max\{0, x\}$. Figure 2.1-(b) presents an example of per-period cost as a function of demand and mode choice. The cost $C^1(k,d)$ depends explicitly on the capacity $k$ of the contract. We assume that the spot market price is larger than the contract price, i.e. $c' > c$. Otherwise, the problem is trivial as the contract mode would never be cost-efficient.

Under stationary demand, when the horizon of available information is more than the commitment period ($\gamma >> \tau$), the problem approaches the full information setting. On the other hand, if the information horizon is significantly smaller than the commitment horizon ($\gamma << \tau$), the problem can be modeled as in the null information case. This chapter considers the scenario where $\gamma < \tau$ but both are of similar magnitude, representative of the scenario faced by the furniture company where $\gamma = 1$ week and $\tau = 4$ weeks.

Table 2.1 presents three settings for the decision process. Given a set of feasible capacities, the decision maker decides whether to sign a commitment contract (and the capacity at which to sign the contract) or use the spot market for the entire next period demand. In Section 2.3.2 we study the setting in which the decision maker has only one choice of capacity with value $\bar{k}$. In Section 2.3.3 we extend the analysis to the case where the capacity can be chosen from a continuum. In the appendix A.3, we briefly discuss the case where capacity can be chosen from a discrete set of capacities. When the decision maker chooses a capacity, we assume he/she chooses the most cost-efficient

| Decision | Type of Contract | | |
| --- | --- | --- | --- |
| | Single Capacity | Continuum Set of Capacities | Discrete Set of Capacities* |
| **Contract mode option** | $m = 0$ or $m = 1$ | $m = 0$ or $m = 1$ | $m = 0$ or $m = 1$ |
| **Capacity Level (if contract is signed)** | Fixed $\bar{k}$ (no decision) | Choose from $k \in [0, \infty)$ | Choose from $k \in \{k_0, k_1, ..., k_M\}$ |

Table 2.1. Contract decision settings.*discussed in the appendix A.3

capacity among the set of capacities available. This decision depends on the costs $c$ and $c'$ as well as the distribution of the unobserved demand. More importantly, the decision also depends on the demand visibility available at the moment of signing the contract.

### 2.3.1 Preliminaries

We consider the decision process faced by the furniture company where managers have demand visibility for the next period; i.e. $\gamma = 1$, but the demand of the remaining periods in the planning horizon is stochastic with a stationary distribution $F$. We denote as $D^w$ the demand when it is a random variable and $d$ when it is a specific realization of the demand.

Assuming that no contract is signed at the end of period $t$, the decision making sequence is as follows: at the end of period $t$ the decision maker sees demand $d$ for period $t+1$ and must then decide whether to sign the commitment contract ($m = 1$) or not ($m = 0$). With $m = 0$, the same sequence is repeated at the end of time $t + 1$. With $m = 1$, there is no decision to be made until the end of time $t + \tau$ when the system resets. The evolution of the system can be formulated as a discounted infinite time horizon Markov Decision Process (MDP). Consider the decision process in a certain period when the observed demand for the next period is $d$, given a discount factor $\delta$. The value function

$V(d)$ can be obtained as the minimum between the cost of serving all the next period demand with the spot market and the cost of signing the contract.

When the spot market is used to meet the entire transportation demand $d$ for the next period, the total cost is the sum of the spot market cost to serve the next period demand and the discounted expected cost when facing the same decision in the following period ($\mathbb{V} = \mathbb{E}_{D^w}[V(D^w)]$):

$$\text{Total Cost}_{m=0} = C^0(d) + \delta\,\mathbb{V}\ \cdot \tag{2.2}$$

When a contract is signed, the capacity level $k$ must also be chosen (unless the set of capacity consists of a single capacity option). If capacity level can be chosen, the decision maker can select capacity based on the demand observed $d$. We make this dependence of the capacity on the observed demand explicit by denoting the selected capacity by $k(d)$. The total cost, when $m = 1$, is the sum of three components:

1. The cost incurred in the following period using the contract with the observed demand $d$, $C^1(k(d), d)$.
2. The discounted expected cost of the following $\tau - 1$ periods in the commitment horizon
   ($\sum_{i=1}^{\tau-1} \delta^i\, \mathbb{E}_{D^w}[C^1(k(d), D^w)]$).

   As the demand distribution is stationary, the expected cost is the same in all periods. Additionally, note that $\sum_{i=1}^{\tau-1} \delta^i = \delta\frac{1-\delta^{\tau-1}}{1-\delta}$. We denote this term as the *discounted commitment horizon* $\Delta(\tau) = \delta\frac{1-\delta^{\tau-1}}{1-\delta}$, which is always bounded between 0 (when $\delta = 0$) and $\tau$ (when $\delta = 1$). Hence, we can write this component as follows:

$$\sum_{i=1}^{\tau-1} \delta^i \, \mathbb{E}_{D^w}[C^1(k(d), D^w)] = \Delta(\tau) \, \mathbb{E}_{D^w}[C^1(k(d), D^w)]$$

3. The discounted expected cost of the decision to be made after the commitment contract expires in $\tau$ periods ($\delta^\tau \, \mathbb{V}$)

The sum of the first and the second components is the total cost incurred during the commitment horizon and is denoted as $C^{1,\tau}(k(d), d) = C^1(k(d), d) + \delta \frac{1-\delta^{\tau-1}}{1-\delta} \, \mathbb{E}_{D^w}[C^1(k(d), D^w)]$. Combining the three terms, we can write the total cost when $m = 1$ as follows:

$$\text{Total Cost}_{m=1} = C^{1,\tau}(k(d), d) + \delta^\tau \, \mathbb{V}\,. \tag{2.3}$$

The value function (i.e. the cost of the decision) can be expressed by the Bellman Equation (using Equations (2.2) and (2.3)) as follows:

$$\begin{aligned} V(d) &= \min\{\text{Total Cost}_{m=0}, \text{Total Cost}_{m=1}\} \\ &= \min\{C^0(d) + \delta \, \mathbb{V}, C^{1,\tau}(k(d), d) + \delta^\tau \, \mathbb{V}\}. \end{aligned} \tag{2.4}$$

At each decision epoch, the decision maker considers the observed demand in the next period, $d$, and commits to a capacity level $k(d)$ from the choice set. We next show that under certain conditions we can restrict attention to functions $k(d)$ that make the decision monotonically increasing with the demand observed; i.e., if for some observed demand it is optimal to commit, then it is also optimal to commit for larger demand. Similarly, if for some observed next period demand it is optimal to entirely use the spot market without commitment, then this option is also optimal for lower values of observed demand. The following lemma characterizes conditions under which the policy obtained is monotonically increasing.

**Lemma 1.** *Consider the cost incurred over the commitment horizon:* $C^{1,\tau}(k(d), d)$*. There exists a monotonically increasing policy for* $V(d)$ *that is optimal if the contract option is such that:*

$$\frac{d(C^{1,\tau}(k(x),x)}{dx} \leq c' \quad \forall x \in \{x | C^{1,\tau}(k(x), x) \text{ is differentiable } \}. \tag{2.5}$$

*In such case, we say that the contract and the selection of the capacity are Efficient.*

We prove the results in the appendix A.1. The condition in Lemma 1 states that a contract and the selection of the capacity are *Efficient* when the cost per unit of observed demand is at most the spot market cost $c'$. Lemma 1 allows us to explore the structure of the solution and obtain intuition on the role of partial information and the trade-off between matching costs to observed demand and factoring costs throughout the commitment horizon. The condition specified in Lemma 1 is found to naturally hold when $c' > c$ as shown in the following Theorem.

**Theorem 1** (Generalization of Lemma 1)**.** *If* $c' > c$*, there exists a monotonically increasing policy for* $V(d)$ *that is optimal.*

Therefore, the condition of being Efficient in Lemma 1 is a direct consequence of $c' > c$.

### 2.3.2 Commitment to a Contract with a Single Capacity Option

Consider the case where the decision maker commits to a contract with a single capacity option $\bar{k}$. At time $t$, the decision maker observes demand $d$ for period $t+1$ and must decide the best of two options: sign a committed contract with capacity $\bar{k}$ for the next $\tau$ periods; or serve all demand for period $t+1$ with the spot market at cost $c'd$. This decision is impacted by the demand observed for period $t+1$. If the demand observed for period $t+1$ is sufficiently low (say close to 0), the decision maker may be better off by meeting the needs of period $t+1$ with the spot market and delaying the decision of signing a contract as committing to capacity $\bar{k}$ will result in most of the capacity being wasted for period $t+1$.

We show that the availability of partial demand information results in outcomes that are fundamentally different than the null information case. Under the case of null information, the commitment time $\tau$ is irrelevant and the contract is signed only if the expected single period cost from committing to capacity $\bar{k}$ ($\mathbb{E}_{D^w}[C^1(\bar{k}, D^w)]$) is no larger than the expected cost of the spot market ($\mathbb{E}_{D^w}[C^0(D^w)]$). Thus, when information is not available, a commitment contract is never signed if :

$$\mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] - \mathbb{E}_{D^w}[C^0(D^w)] > 0, \tag{2.6}$$

However, in the partial information case a commitment contract can be signed even if the condition in equation (2.6) is true. We next provide a condition when the commitment contract is not signed, regardless of visible demand.

**Proposition 1.** *Assume that the decision maker can see the next period demand. It is never optimal to sign the commitment contract if:*

$$\mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] - \mathbb{E}_{D^w}[C^0(D^w)] > \frac{\bar{k}(c' - c)}{\Delta(\tau)}. \tag{2.7}$$

**Definition 1.** *If:*

$$\mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] - \mathbb{E}_{D^w}[C^0(D^w)] \leq \frac{\bar{k}(c' - c)}{\Delta(\tau)},$$

*we say that a Cost Commitment Balance Condition (CCBC) holds.*

We use the term *Cost Commitment Balance Condition* to highlight the relationship between the one-period cost difference of the two modes and the cost of commitment. The left hand side of the CCBC can be interpreted as the difference in expected per period cost of the two modes when the available partial information is ignored. The right hand side represents the maximum benefit that commitment can provide in a period, prorated by the discounted commitment horizon.

The right hand side of the CCBC shows that if the left hand side is positive (condition in eq. (2.6) true) when the commitment horizon $\tau$ is large enough (and $\delta$ close enough to 1) the CCBC will not hold and, therefore, the commitment contract should never be signed. As expected, as the commitment horizon increases, the commitment contract becomes less attractive.

From Proposition 1 and a comparison of equations (2.6) and (2.7), one can see that the availability of partial information reduces the instances where no commitment contract is signed because $\frac{\bar{k}(c'-c)}{\Delta(\tau)} > 0$. If a decision maker knows that the demand in the next period is high, he/she may be willing to commit to a contract that would not be signed in the absence of information. Thus, commitment contracts that would not be cost effective in the null information case can become cost effective when partial information is available.

Proposition 2 identifies the optimal policy for the decision maker when the CCBC holds:

**Proposition 2.** *Assume that the decision maker can see the next period demand $d$. If the CCBC holds, there exists a unique threshold policy with threshold $\hat{d} < \bar{k}$ such that: If the next period demand is $d > \hat{d}$, it is optimal to sign the commitment contract. If $d \leq \hat{d}$, it is optimal not to sign the contract and use the spot market for all the demand in the next period, delaying the commitment decision for one more period. The threshold $\hat{d}$ is the unique solution of the following equation:*

$$\mathbb{C}^0(x) - \mathbb{C}^1(\bar{k}, x) + (1 - F(x) + \tfrac{1}{\Delta(\tau)})(C^0(x) - C^1(\bar{k}, x)) = 0,$$

*where:*

$$\begin{aligned} \mathbb{C}^0(x) &= \textstyle\int_{d \leq x} C^0(d) dF(d) \\ \mathbb{C}^1(x) &= \textstyle\int_{d \leq x} C^1(c, d) dF(d). \end{aligned} \tag{2.8}$$

The equation in Proposition (2) is obtained by assuming that both modes are equally efficient when $d = \hat{d}$ and using the information to solve the Bellman Equation.

Proposition 2 shows that under CCBC, the decision to commit to a contract explicitly depends on the value of the observed demand in the next period. If the observed demand is sufficiently high, the decision maker is willing to commit, but not otherwise. It is this partial visibility that results in the decision maker accepting contracts that would have been rejected in the null information case.

The availability of partial information also results in instances where the decision maker does not commit to a contract under conditions where he/she would have committed in the absence of information. This happens when the visible demand for the next period is low ($d < \hat{d}$). In such a situation, the decision maker is better off by meeting all transportation needs for the next period from the spot market ($m = 0$). Thus, the availability of partial information may result in fundamentally different decisions from the null information case.

In Figure 2.2 we use a numerical example to compare the policies in the null information and partial information cases. Due to its generality, we use the gamma distribution for the demand. The parameters scale ($\theta$) and shape ($a$) accordingly to set the expected value and the variance as described in the legends. The demand follows gamma distribution with expected value 0.66 and standard deviation 0.66. We increase the cost of contracted capacity from 0 to 1 and hold the remaining parameters constant: $\tau = 4$, $c' = 1$, $\bar{k} = 1$ and $\delta = 0.9$. Experiments with other distributions can be found in the appendix (e.g. Figure A.3-Figure A.7 in appendix A.5), which show that, qualitatively, results hold for other distributions.

Figure 2.2-(a) presents the optimal policy for the null information case as a function of the ratio of $c$ and $c'$. Given the fixed parameters, if $c/c' < 0.51$, the contract is always signed. It is never signed otherwise.

In Figure 2.2-(b) we consider the partial information case and illustrate how the threshold $\hat{d}$ changes with $c$. The y-axis represents the demand observed for the next period. The

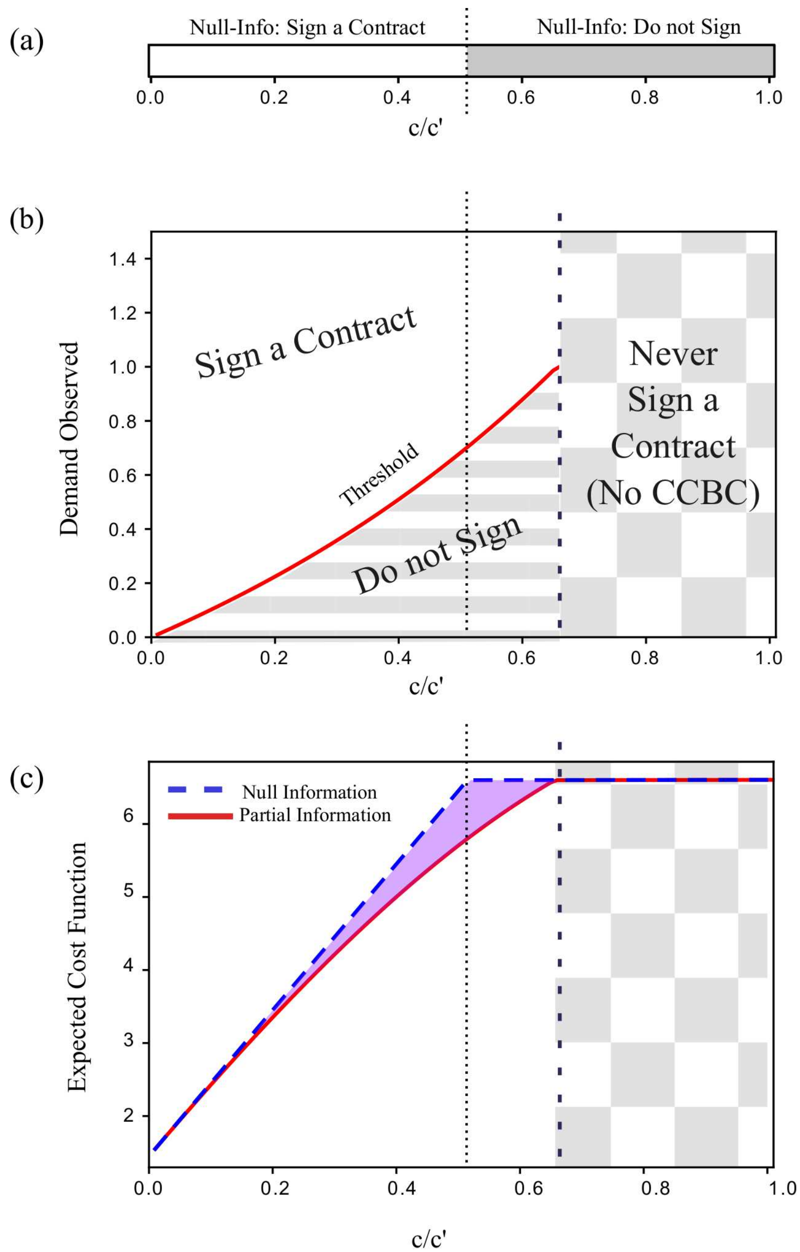


Figure 2.2. Contract policies as a function of contract and spot market costs ($c/c'$). (a): Null information policy. (b): Partial information policy (c): Expected cost of partial and null information policies.

dashed line separates the region into CCBC (left) and no-CCBC (right) (at $c/c' = 0.66$). When $c/c' > 0.66$, the contract is never signed regardless of the value of the observed next period demand. Note that for the same demand distribution and capacity level, this cut-off point is higher than the commitment cut-off point (0.51) in the null information case. When $c/c' < 0.66$ (CCBC holds), the threshold (solid line) separates the regions where the decision maker commits or does not commit based on the observed next period demand. When the demand observed is above the solid line, the commitment contract is signed. Otherwise, the spot market is used for all the transportation needs for the next period.

We compare the expected value function (the total expected cost incurred) of the partial and null information scenarios in Figure 2.2-(c). When CCBC is satisfied ($c/c' < 0.66$), the expected cost of the partial information case is always lower. When $c/c'$ reaches the region where the CCBC does not hold, the decision maker never commits and, therefore, both the null information and the partial information have the same costs. The partial information is most valuable around the $c/c'$ threshold for the null information case ($c/c'$ $= 0.5$ in our numerical example).

**The Value of Information with a Single Capacity Option.** We study the effect of demand volume on the value of information (Figure 2.3). For this purpose, we keep the Coefficient of Variance (CV) constant while changing the expectation of the demand ($\mathbb{E}$[D]). Additional experiments (Figure A.3 in appendix A.5) confirm that these insights hold for other distributions and with varying CV.

We first consider the length of the information horizon relative to the commitment horizon. If the commitment horizon $\tau$ is much longer than $\gamma = 1$, the information provided by the next period demand is small compared to the uncertain demand that is not visible. In this case, partial information does not add significant value as the cost incurred under partial information is nearly the cost of having no information. On the other hand, if the commitment horizon is close to the information horizon ($\tau$ approaches 1), the visibility of next period demand is likely to be quite valuable. This observation is validated by our numerical experiments shown in Figure 2.3, where we show the evolution of the relative difference of the expected costs of the partial information and the null information scenarios. When the information of the next period demand is valuable, the cost incurred under the partial information is significantly smaller than the cost incurred under null information. In all experiments, we fix $\bar{k} = 1$, $c/c' = 0.5$ and $\delta = 0.9$. Demand follows a gamma distribution with an expected value that ranges from 0.1 to 1.5 and a fixed CV of 0.7. In Figure 2.3-(a) and 2.3-(b) we increase the ratio $1/\tau$, that represents the fraction of commitment horizon that is visible when the decision is made. We show the change of the relative difference of costs for different expectations of the demand ranging from 0.1 to 1.5. In Figure 2.3-(a) we show the results for different expectations ranging from 0.1 to 0.7 in (a) and in Figure 2.3-(b) we show the results for expectations ranging from 0.7 to 1.5. The higher is the color intensity of each curve, the higher is the demand expectation. Observe that the value of partial information increases as $1/\tau$ increases (this happens when the commitment horizon $\tau$ decreases). As expected, the value of partial information increases as the information horizon represents a greater fraction of the commitment horizon.

If the expected value of demand is much lower than the capacity of the commitment contract, commitment is unlikely because the observed demand for the next period is likely to be lower than the threshold $\hat{d}$. In such a situation, partial information is unlikely to be useful because it will not change the commitment decision. When the expected value of demand is much larger than the commitment capacity, the decision maker will almost

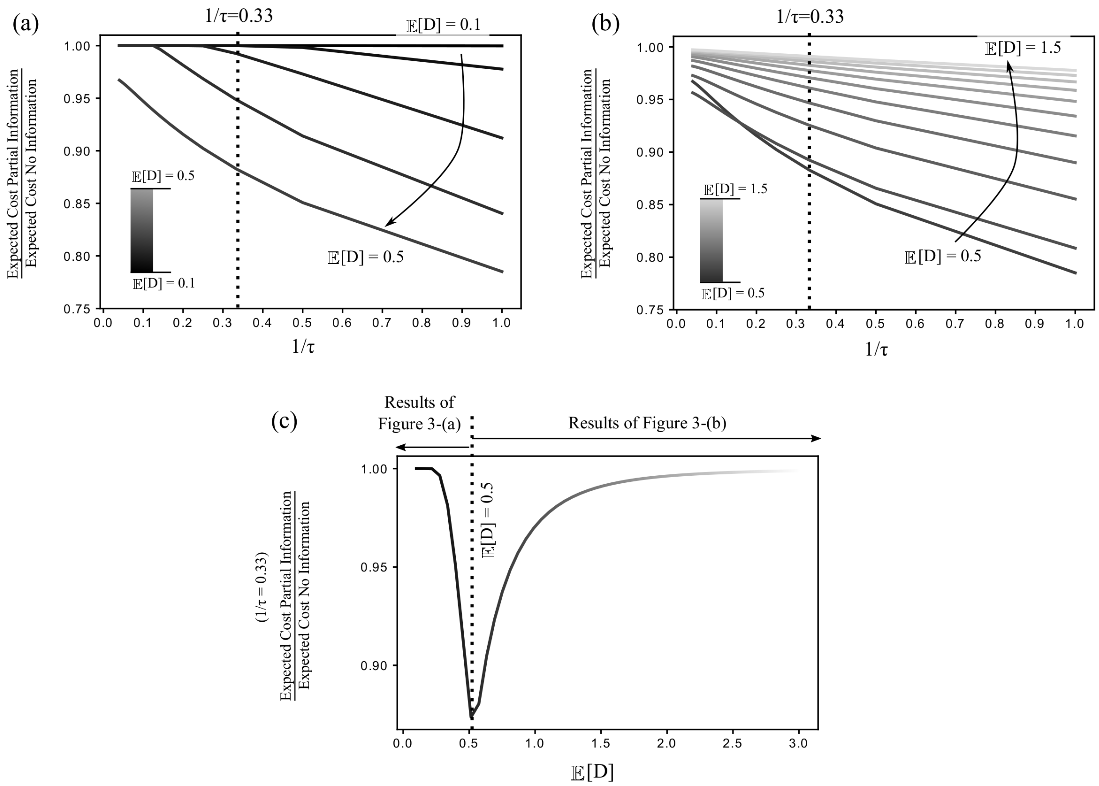


Figure 2.3. (a): The value of one period demand visibility with low expected demand, $\mathbb{E}[\mathrm{D}]$. (b): The value of one period demand visibility with high expected demand (c): The value of one period demand visibility as a function of expected demand.

certainly commit because the observed demand is likely to be larger than the threshold $\hat{d}$. In this case, partial information is again unlikely to be useful. Thus, partial information is most useful when the expected value of demand per period is close to the commitment capacity $\bar{k}$. This observation is validated in Figure 2.3-(c), where we fix $1/\tau = 0.33$ and we change the expectation of demand from 0.1 to 3. The value of partial information diminishes when expected demand is much lower or much higher than $\bar{k} = 1$. For very small or very large values of expected demand, the ratio of expected cost in the partial and null information cases converges to 1. It is for intermediate values of expected demand where partial information provides the highest benefits, allowing 15% of cost reduction when the expected demand is 50% percent of the commitment capacity.

**The Impact of Commitment and Size of the Lane on the Threshold Policy.** We explore the effect of the commitment horizon and the expected demand on the threshold policy in Figure 2.4 where we show the threshold when the ratio $1/\tau$ and the expected demand change while the CV is constant at 0.7[1].

The threshold decreases as the expected demand increases. Intuitively, when the expected demand is low relative to capacity, using the spot market is likely to be more cost effective than commitment. Thus, the threshold is high because observed demand in the next period must be quite high to make the contract cost-efficient over the entirety of the commitment horizon. On the other hand, when the expected demand is large relative to capacity, using the spot market is more cost-efficient only if the next period demand is very small, and the threshold is low.

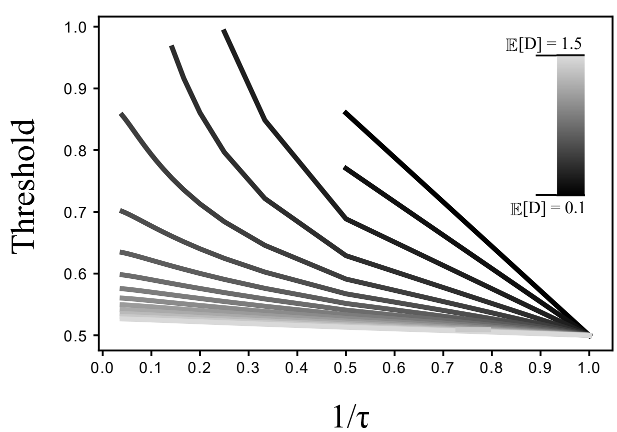


Figure 2.4. Shape of the threshold $\hat{d}$ as a function of expected demand and $\frac{1}{\tau}$.

[1]See Figure A.4 in appendix A.5 for further experiments varying CV.

Furthermore, observe that the threshold decreases as $1/\tau$ increases (as the commitment horizon decreases). When the commitment horizon $\tau$ is long, the commitment contract is less attractive and is signed only if the next period demand is sufficiently large. Interestingly, for the light curves, representing high expected demand, there exists a positive value of the threshold even if $\tau \to \infty$. When $\tau$ is large, the optimal policy becomes a stopping rule policy: The contract is not signed until the next period demand is larger than the threshold and, after this event the contract is used for a long time period. Of course, for such long time commitments, the savings from following this policy become negligible compared to the total cost of the commitment time. Hence, as shown in Figure 2.3, the costs of the null and full information are close when $\tau$ is very large.

Returning to the furniture company, partial information is likely to be valuable because the information horizon of one week represents a meaningful portion of the commitment horizon of four weeks. The partial information is likely to be most valuable along lanes where the quantity transported is within a reasonable range of the available capacity (e.g. a truckload).

### 2.3.3 Commitment to a Contract from Among a Continuum Set of Capacities

We extend the analysis in Section 2.3.2 to the case where the decision maker has a continuum of capacity choices available, each time he/she decides to commit. The capacity $k_v$ selected is a decision variable that can be changed each time the decision maker signs a new commitment contract. After selecting $k_v$, the contract has capacity $k_v$ for the next $\tau$ periods. Once the commitment expires, a new value of $k_v$ can be selected for the next commitment contract. We call $C^1(k_v, d)$ the per-period cost of the contract option as a function of the demand ($d$) and the capacity chosen ($k_v$). Equation (2.9) shows the cost structure of each mode.

$$
\begin{aligned}
&C^0(d) = c'd && \text{for } m = 0 \\
&C^1(k_v, d) = ck_v + c'[d - k_v]^+ && \text{for } m = 1.
\end{aligned} \tag{2.9}
$$

Using $\mathbb{E}_{D^w}[C^1(k_v, D^w)]$ and $\Delta(\tau)$ as defined in equations (2.2) and (2.3), the expected cost of committing to a capacity $k_v$ is denoted by $C^{1,\tau}(k_v, d)$:

$$
C^{1,\tau}(k_v, d) = \; C^1(k_v, d) + \Delta(\tau)\, \mathbb{E}_{D^w}[C^1(k_v, D^w)]. \tag{2.10}
$$

We denote by $k^*(d)$ the optimal capacity at which the decision maker signs the commitment contract when the next period demand is known to be $d$.

We now show that partial information results in fundamentally different policies by the decision maker relative to the null information scenario. When no information is available, the commitment horizon is irrelevant: the decision maker always signs a commitment contract at a capacity level $k^*_{null}$, where the selected capacity is obtained by solving a simple newsvendor problem as follows: $k^*_{null} = F^{-1}(\frac{c'-c}{c'})$. We show that in the partial information scenario, the decision maker may select different levels of capacity each time he/she commits, based on the observed demand.

**Proposition 3.** *Assume that the decision maker sees the next period demand $d$. If the decision maker signs a commitment contract, the optimal capacity signed, $k^*(d)$, is completely determined by $d$ as follows: Define $k^*_l := F^{-1}(\frac{c'-c}{c'} - \frac{c}{c'\Delta(\tau)})$ and $k^*_u :=$*

$F^{-1}(\frac{c'-c}{c'} - \frac{c}{c'\Delta(\tau)} + \frac{1}{\Delta(\tau)})$. *Then,*

$$k^*(d) = \begin{cases} k_l^* & \text{if } d < k_l^* \\ d & \text{if } k_u^* \geq d \geq k_l^* \\ k_u^* & \text{if } d > k_u^*. \end{cases}$$

We remark on some interesting properties of the selected capacity $k^*(d)$:

(1) In contrast to the null information case where the same level of capacity is selected for each commitment horizon, under partial information the selected capacity may vary with the observed demand for the next period. The structure of $k^*(d)$, shown in Proposition 3, is such that if the observed demand $d$ is in the set $UC = [k_l^*, k_u^*]$ then the signed capacity is exactly equal to the demand, i.e. $k^*(d) = d$. If the observed next period demand $d$ is larger than $k_u^*$ (hence $d \notin UC$) the signed capacity is $k_u^*$. Similarly, if the observed next period demand $d$ is lower than $k_l^*$ (hence $d \notin UC$) the signed capacity is $k_l^*$. Hence, under partial information, the chosen capacity explicitly depends on the demand observed.

(2) The visibility of next period demand $d$ allows the decision maker to select a capacity level between $d$ (a myopic choice to minimize cost for the next period) and $k_{null}^*$ (an expected-cost choice to minimize cost for the rest of the commitment horizon). Observe that $k_{null}^*$ is always in $UC$: $k_l^* < k_{null}^* < k_u^*$. The band of $UC$ can be interpreted as the range of adaptation the decision maker has when selecting the contracted capacity.

(3) As shown in Figure 2.5-(a), the width of $UC$ depends only on the distribution of the demand ($F$) and the discounted commitment horizon ($\Delta(\tau)$). As the length of the commitment horizon increases, the width of $UC$ shrinks. When $\tau >> \gamma$, the $UC$ consists of a single element $UC = \{k^*_{null}\}$, and we move from the myopic choice to the expected-cost choice.

We now show that the partial visibility of demand also impacts whether the decision maker should commit or not. Whereas the decision maker always commits in the null information scenario when a continuum of capacities is available, the decision maker may choose not to commit under certain values of observed demand and delay a capacity choice until demand increases. We first compare the costs of committing and not committing when next period demand $d$ is visible. By substituting the expression of $k^*(d)$ in the second equation of (2.9) we obtain that the first period cost if we commit at capacity $k^*(d)$ ($C^1(k^*(d), d)$) is a piece-wise linear convex function that depends on the observed demand $d$ as shown in equation (2.11) and Figure 2.5-(b)).

$$C^1(k^*(d), d) = \quad ck^*(d) + c'[d - k^*(d)]^+ = \begin{cases} ck^*_l & \text{if } d < k^*_l \\ cd & \text{if } k^*_u \geq d \geq k^*_l \\ c'd - (c' - c)k^*_u & \text{if } d > k^*_u. \end{cases} \tag{2.11}$$

Similar to equation (2.4), let $V(d)$ be the value function after seeing next period demand $d$ and $\beta_f(d)$ the expected value of $C^1(k^*(d), D^w)$ over $D^w$. We formulate the evolution of the system as follows:

$$V(d) = \min\{C^0(d) + \delta\,\mathbb{V}, C^1(k^*(d), d) + \Delta(\tau)\beta_f(d) + \delta^\tau\,\mathbb{V}\}. \tag{2.12}$$

Next we show that the flexibility for the decision maker to select different levels of capacity for each commitment contract ensures that the CCBC condition always holds at the optimal capacity level. This is in contrast to the previous section where the CCBC condition did not always hold when the decision maker could only commit to a single capacity.

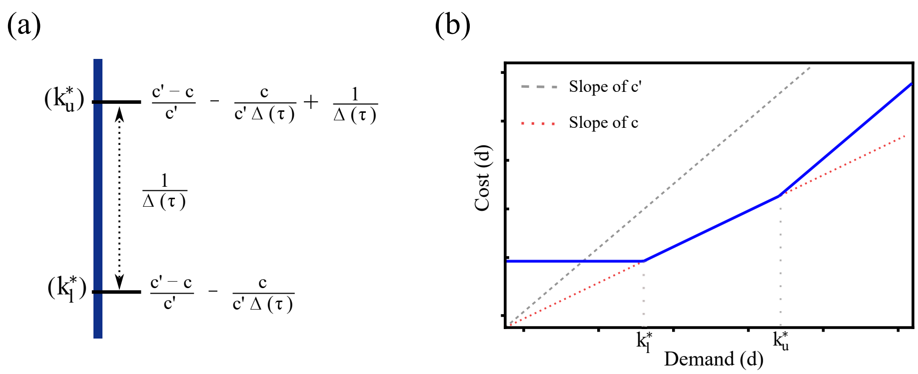


Figure 2.5. (a): Range of Useful Capacities (argument of $F^{-1}$). (b): Cost of the first period: $C^1(k^*(d), d)$.

**Lemma 2.** *Assume that the decision maker can see next period demand. Under commitment to a contract from among a continuum set of capacities, the following is always true:*

$$\mathbb{E}_{D^w}[C^f(k_u^*, D^w)] - \mathbb{E}_{D^w}[C^0(D^w)] < \frac{k_u^*(c' - c)}{\Delta(\tau)}. \tag{2.13}$$

*As a result the Continuous Cost Commitment Balance Condition (C-CCBC) always holds at the optimal level of signed capacity.*

Note that the left hand side of the C-CCBC depends on $k_u^*$ which depends on the discounted commitment horizon and the distribution of the demand.

The expected one period cost of using the spot market is equal to the expected one period cost of committing at capacity $k_v = 0$. Also, $k_{null}^*$ minimizes the expected one period cost of the contract option. Therefore, $\mathbb{E}_{D^w}[C^f(k_{null}^*, D^w)] - \mathbb{E}_{D^w}[C^0(D^w)] < 0$. Consequently, for the case of null information, where the commitment horizon becomes irrelevant and the capacity chosen is always $k_{null}^*$, a contract with capacity $k_{null}^*$ is always signed. In contrast, when partial information is available the next result shows that the decision maker may or may not commit depending on the value of the observed next period demand $d$.

**Proposition 4.** *Assume that the decision maker sees the next period demand $d$. There always exists a threshold policy with threshold $\hat{d} < k_u^*$ such that: If the next period demand is $d \geq \hat{d}$, it is optimal to sign the committing contract with capacity $k^*(d)$ (eq. 2.11). If $d \leq \hat{d}$, it is optimal to use the spot market for all the demand in the next period, delaying the commitment decision for one more period. The threshold $\hat{d}$ is the unique solution of the following equation:*

$$\begin{aligned} 0 = \mathbb{C}^0(x) + \int_x^\infty C^1(k^*(d), d)dF(d) + \Delta(\tau)\int_x^\infty \beta_f(d)dF(d) + \\ +(1 - F(x) + \tfrac{1}{\Delta(\tau)})(C^0(x) - C^1(k^*(x), x) - \Delta(\tau)\beta_f(x)). \end{aligned} \tag{2.14}$$

The result above shows that under partial information, a contract is signed if the demand of the next period is sufficiently high. When the observed next period demand is low, the optimal decision is to use the spot market for all the observed demand. Such a policy is fundamentally different from the null information scenario where a commitment contract is always signed. The availability of the spot market at cost $c'$ allows the decision maker to delay the commitment decision. The decision maker is willing to spend a higher cost of $c'$ per unit for the low next period demand and wait for a period with higher demand before committing. Thus, the availability of partial information allows the spot market to be used as an option to acquire more information (wait a period) if the currently available information does not favor commitment (visible demand is low).

To summarize, when the decision maker is allowed to choose the capacity from a continuum, the decision of signing the contract and its level of capacity depends on the information available. When he/she has no demand information, the contract is always signed at a fixed capacity level. In contrast, when the demand information is available for the next period, the contract is signed only when the demand is known to be sufficiently high. Not only does the decision of signing the contract explicitly depend on the demand observed, but the level of capacity at which the decision maker commits is itself a function of the visible next period demand.

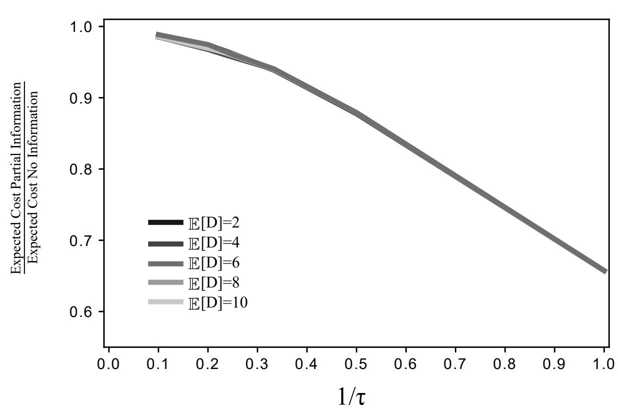


Figure 2.6. Value of information when decision maker chooses the capacity level.

**The Value of Information with a Continuum Set of Capacity Options.** We present numerical experiments to illustrate the value of information as expected demand changes (Figure 2.6). In the experiments, we use $\delta = 0.9$ and assume that demand follows a gamma distribution with a constant CV of 0.7[2] and an expected value increasing from 2 to 10 represented in dark and light curves, respectively. We observe in Figure 2.6 that the relative difference of the expected cost incurred under partial and no information remains the same when the expected value of the demand is changed (the dark solid line is the result of the 5 overlapped identical lines). Our observation indicates that if the CV is kept constant the problem scales with the expected demand and, therefore, the threshold should increase with expected demand.

**The Impact of Commitment, Costs and Size of the Lane on the Threshold Policy.** To understand why the value of information is the same for any size of the demand we focus on how the threshold depends on the demand. In Figure 2.7-(a) we fix the cost ratio to $c/c' = 0.5$ and show the threshold for different values of $1/\tau$ and expected demand (gray scale: the lighter the line the higher is the expected demand) while the CV is kept constant to 0.7[3]. The threshold increases with the expected value of demand unlike the results obtained with a single capacity set (Section 2.3.2 and Figure 2.4), where the threshold decreases with the expected demand. When the decision maker can choose the capacity, the chosen capacity level increases with the expected demand (i.e. in Proposition 3, $k_l^*$ depends on the inverse of the distribution $F$). Thus, the contract is signed only when the observed demand is sufficiently high to make the contract, at capacity level higher or equal than $k_l^*$, cost-efficient.

---

[2]See Figure A.6 in appendix A.5 for further experiments varying CV.
[3]See Figure A.7 in appendix A.5 for further experiments varying CV.

As noted earlier, when the commitment horizon is large ($\tau \to \infty$), the threshold is positive and, hence, the optimal policy becomes a stopping rule: as shown in Figure 2.7-(a) the threshold is high and the spot market is used until the observed next period demand is higher than the threshold. At that point, the decision maker commits with a capacity level $k^*(d) = k_{null}$.

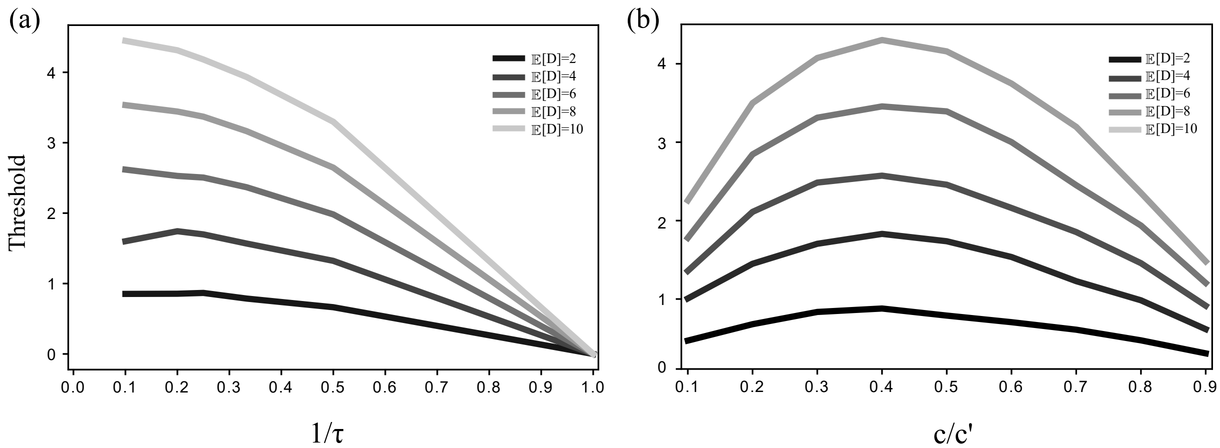


Figure 2.7. Numerical results for the continuum set of capacities case. (a): Threshold curves changing with commitment horizon. (b): Threshold curve changing with the ratio $c/c'$.

We also examine how the spot market and contract costs change the optimal policy. The contract cost $c$ determines how the modes are used: When the cost $c$ is low relative to $c'$, the decision maker has low commitment risk and always signs the contract with high capacity. When the cost $c$ is high (close to the spot market $c'$), the decision maker commits to a low capacity to serve the baseline demand, and generously uses the spot market to serve high demand fluctuations. When the cost is between these two extremes, the spot market and contract are both used to serve fluctuations and baseline demand. We validate these observations numerically in Figure 2.7-(b). In the experiments we fix the commitment horizon to $\tau = 4$ ($1/\tau = 0.25$). We show how the threshold changes when we modify the fraction $c/c'$ and expected demand (gray scale). When the fraction $c/c'$ is close to 0 or 1, the optimal threshold is low. When $c/c'$ is between 0.2 and 0.7 the optimal threshold is higher, hence only when the next period demand is sufficiently high, the contract is signed. Interestingly, the contract is signed with high probability (threshold set low) in both extremes values of the cost fraction $c/c'$. This result contrasts with the conclusions in Section 2.3.2, where only one capacity is available and the threshold is always increasing with relative cost $c/c'$.

In the case of the furniture company, there is clear value from partial information when a range of capacity levels exist for contracts. Any such flexibility from the carriers would allow the furniture company to properly choose capacity levels based on observed demands.

### 2.3.4 Seasonal Demand

Often, demand on transportation lanes exhibits seasonality, as observed by the furniture company studied in this work. Figure 2.8 shows average daily transportation needs of the company in three lanes for the 2016-2017 fiscal year. These lanes correspond to three hub-spoke pairs with the largest demand volume of a region that contains a total of 20 lanes used by the company. Demand units are scaled for confidentiality but proportional across periods and lanes. The most pronounced peaks in the season occur in the weeks during and after special dates (Black Friday, President's Day and Labor Day) where the volume of orders is known to be high. These peaks are present on the same dates in all lanes and the same pattern repeats from one year to another. Transportation procurement decisions are made in the framework of this seasonal demand.

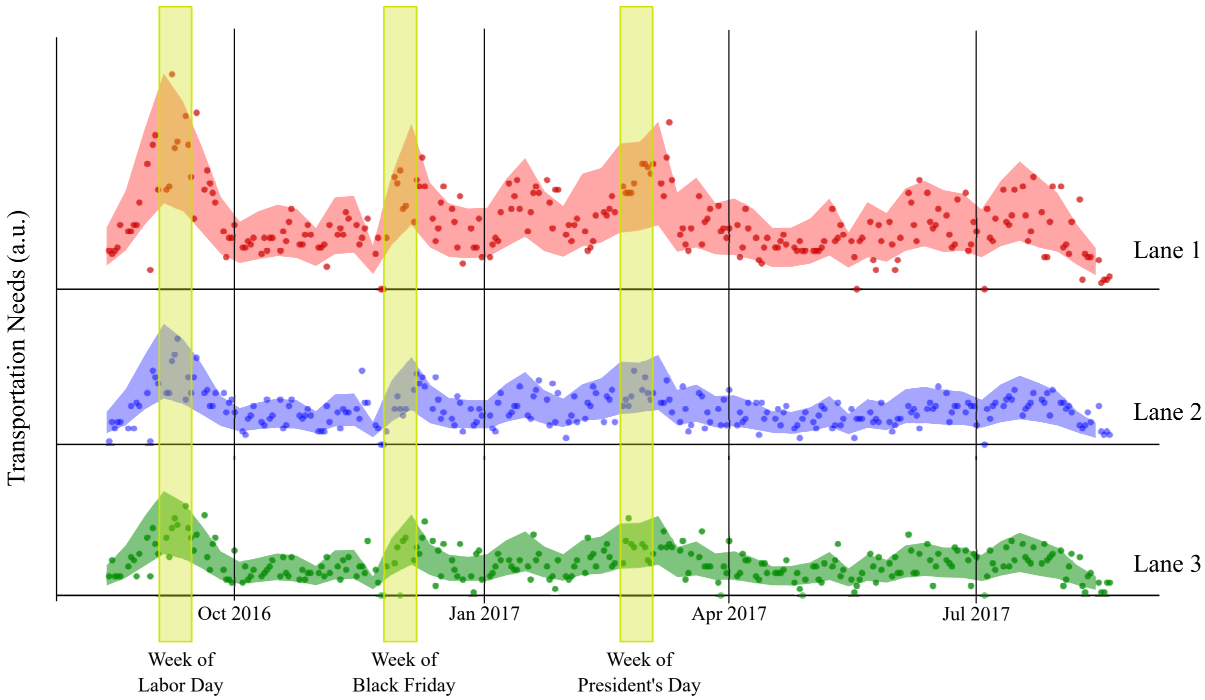


Figure 2.8. Transportation needs for three lanes (units scaled) for the year 2016-2017. Dots show daily transportation needs. Shade shows the estimated mean ± estimated standard deviation.

We model the case with stationary but seasonal demand, when capacity of the contract option can be chosen from a continuum as in Section 2.3.3. We consider a year that consists of $0, \ldots, T-1$ periods. The distribution of the demand of all periods is known, denoted by $F_t$ with $t = 0, \ldots, T-1$, and we assume that $F_{t+T} = F_t$ for all $t = 0, 1, 2, \ldots$.

First, we focus on the capacity decision assuming a contract is signed. With capacity $k_v$, the cost over the commitment horizon can be calculated as the sum of the cost of the period where the demand realization is known and the expected cost of the following $\tau - 1$ periods. Let $\mathbb{C}^1_{t'}(k_v) = \int (ck_v + c'[d' - k_v]^+) dF_{t'}(d')$ be the expected cost of period $t'$. When a contract is signed at the end of period $t-1$ and the next period demand realization is known to be $d_t$, the cost during the commitment horizon is:

$$C_t^{1,\tau}(k_v, d_t) = \quad ck_v + c'[d_t - k_v]^+ + \textstyle\sum_{i=1}^{\tau-1} \delta^i \mathbb{C}^1_{t+i}(k_v). \tag{2.15}$$

Equation (2.15) has the same structure as Equation (2.10) and it is easy to see that the capacity that minimizes $C_t^{1,\tau}(k_v, d_t)$ has the same structure as the solution in Proposition 3 with upper and lower thresholds that depend on the period, denoted $k^*_{u,t}$ and $k^*_{l,t}$ for $t = 0, 1, \ldots, T$.

**Proposition 5.** *Assume that demand is seasonal over $T$ periods and the decision maker observes next period demand $d_t$. If the optimal decision at period $t$ is to sign the contract, the capacity signed at time $t$, $k^*_t(d_t)$, is completely determined by the next period demand as follows:*

$$k^*_t(d_t) = \begin{cases} k^*_{l,t} & \textit{if } d_t < k^*_{l,t} \\ d_t & \textit{if } k^*_{u,t} \geq d \geq k^*_{l,t} \\ k^*_{u,t} & \textit{if } d_t > k^*_{u,t} \end{cases},$$

*where $k^*_{l,t}$ and $k^*_{l,t}$ are period-dependent upper and lower thresholds determined as:*

$$k^*_{u,t} = \arg\min_{k_v} ck_v + \textstyle\sum_{i=1}^{\tau-1} \delta^i \mathbb{C}^1_{t+i}(k_v)$$

$$k^*_{l,t} = \arg\min_{k_v} -(c' - c)k_v + c'd + \textstyle\sum_{i=1}^{\tau-1} \delta^i \mathbb{C}^1_{t+i}(k_v) \ .$$

As in Proposition 3, if the observed demand is $d_t \geq k^*_{u,t}$, the optimal capacity signed at time $t$ is $k^*_t = k^*_{u,t}$. If $d_t \leq k^*_{l,t}$, the optimal capacity signed at time $t$ is $k^*_t = k^*_{l,t}$. If $k^*_{u,t} \geq$

$d_t \geq k^*_{l,t}$ the optimal capacity is $k^*_t = d_t$. The capacity signed under null information, $k^*_{null,t}$ always falls between $k^*_{u,t}$ and $k^*_{l,t}$. When signing capacity, the same intuition and results presented for stationary demand hold for seasonal demand (see remarks (1)-(3) in Section 2.3.3).

We now focus on the decision of signing or not signing the contract at period $t$. Let $V_t(d_t)$ be the value function at period $t$ and $\mathbb{V}_t$ its expected value. With seasonal demand, the expected value function at time $t$ and at time $t+T$ is the same ($\mathbb{V}_t = \mathbb{V}_{t+T}$). We have $T$ value functions, each corresponding to making a decision at time $t \in \{0, 1, 2, \ldots, T-1\}$. At each period, the value function consists of the minimum of two: (1) the cost of not signing the contract at time $t$ and revisiting the decision at $t+1$; and (2) the cost of signing the contract with the optimal capacity, committing for $\tau$ periods. The optimality equations that link all the periods can be written as follows for each period:

(2.16)

$$V_t(d_t) \quad = \min\{C^0(d_t) + \delta\mathbb{V}_{t+1}, ck^*_t(d_t) + c'[d_t - k^*_t(d_t)]^+ + \textstyle\sum_{i=1}^{\tau-1} \delta^i \mathbb{C}^1_{t+i}(k^*_t(d_t)) + \delta^\tau \mathbb{V}_{t+\tau}\}.$$

It is easy to see that there exists an optimal policy that is monotonic non-decreasing with the same structure as described in Proposition 4 with period dependent thresholds: At each period $t$, if the demand is larger than a threshold $\hat{d}_t$, it is optimal to sign a contract with capacity $k^*_t(d_t)$. If next period demand is lower than $\hat{d}_t$, it is optimal not to sign any contract. We formalize this statement in Proposition 6.

**Proposition 6.** *Assume that demand is seasonal over $T$ periods and the decision maker sees the next period demand $d_t$. The decision of whether to sign the contract or* *not is completely determined by the next period demand as in Proposition 4 with a period-dependent threshold policy characterized by* $\hat{d}_t$ *for* $t = 0, 1, \ldots, T$.

*The thresholds* $\hat{d}_t$ *uniquely solve the following system of* $2|T|$ *non-linear equations with* $2|T|$ *variables.*

(2.17)

$$C^0(\hat{d}_t) + \delta \mathbb{V}_{t+1} = ck_t^*(\hat{d}_t) + c'[d_t - k_t^*(\hat{d}_t)]^+ + \sum_{i=1}^{\tau-1} \delta^i \mathbb{C}^1_{t+i}(k_t^*(\hat{d}_t)) + \delta^\tau \mathbb{V}_{t+\tau} \quad t = 1, \ldots, T$$

(2.18)

$$\begin{aligned} \mathbb{V}_t \;\; &= \int_{d' < \hat{d}_t)} \Big\{ C^0(d') + \delta \mathbb{V}_{t+1} \Big\} dF_t(d') + \\ &\int_{d' \geq \hat{d}_t)} \Big\{ ck_t^*(d') + c'[d' - k_t^*(d')]^+ + \sum_{i=1}^{\tau-1} \delta^i \mathbb{C}^1_{t+i}(k_t^*(d')) + \delta^\tau \mathbb{V}_{t+\tau} \Big\} dF_t(d') \quad t = 1, \ldots, T. \end{aligned}$$

Comparing this result with those in Section 2.3.3, the only difference is that, under seasonal demand, the threshold depends on the period. All insights and conclusions still apply.

Solving the system for general demand distributions $F_0, \ldots, F_{T-1}$ can be computationally intensive because a single evaluation of Equation (2.18) requires computing a double integral that is moderately sensitive to the approximation error. Instead of directly solving the system of equations, we propose using a sampling framework with the value iteration algorithm, which, for this particular case, can be easily parallelized across periods.

In following sections, we show that when demand is seasonal, partial information can be leveraged by properly coordinating the periods where the lanes are free of commitment with the periods where demand changes sharply.

***Wait-and-See* vs. *Now-or-Never*: Contracting with Seasonal Surges and Drops.** We study how a company can leverage demand information to manage demand surges similar to those in Figure 2.8. For simplicity of exposition, assume that there exists a single demand surge. If the surge is visible, the company can sign a customized capacity contract that adapts to the higher observed demand. The company greatly benefits from demand visibility, the incentives for signing a contract are high and the threshold for committing is low as a result. However, in periods prior to the visibility of the surge, the company is inclined to delay commitment until the demand surge is visible. We confirm this observation in Figure 2.9-(a). (Appendix A.4 includes details about the numerical experiments.) Observe that when $\tau = 3$ the threshold is high in periods 8 and 9 before the surge is revealed in period 10. Signing a contract in either period 8 or 9 would put an established contract in place when the surge is revealed, removing the company's ability to set a customized contract for the surge. As we move farther back from the surge, we observe an oscillating pattern in threshold levels that dampens over time. In period 7, the threshold drops: a contract signed in this period would expire before the surge demand is revealed in period 10 (with $\tau = 3$). The threshold rises again in period 6 as a contract signed in this period removes the ability to sign a contract in period 7, which we just observed is a desirable time to sign a contract since it is offset by $\tau$ periods from the revelation of the surge demand.

We refer to the strategy of coordinating contract commitment in the periods leading up to the reveal of the surge demand as a *wait-and-see* strategy. The oscillating pattern of wait and see is determined such that it is unlikely to have a contract in place when demand for the surge is revealed, allowing the decision maker to make the appropriate capacity choice at that time.

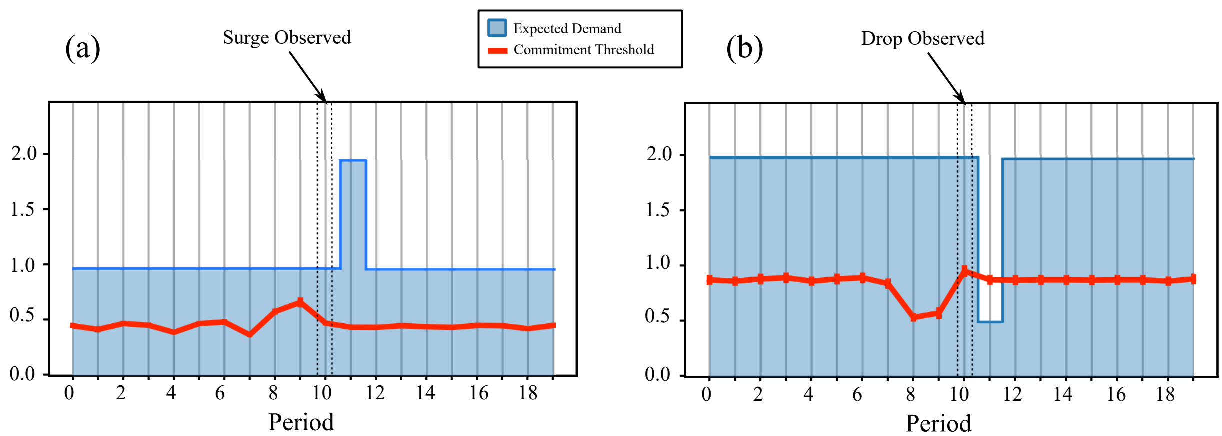


Figure 2.9. Threshold policy with seasonal surge (a) and drop (b) with $\tau = 3$.

We show that a different strategy is optimal when, instead of a surge, a demand drop is expected, as shown in Figure 2.9-(b). It is unlikely that observing demand for the drop would result in signing a contract since demand is likely to be low and the spot market will be more cost-efficient. In periods prior to the visibility of the drop, the company is more likely to take advantage of demand visibility by signing a contract that adapts well to the observed demand in the periods with moderate expected demand. We confirm this observation in Figure 2.9-(b). Observe that when $\tau = 3$ the threshold is low in periods 8 and 9 (higher incentives to sign a contract) because signing a contract implies entering the drop with an established contract. Therefore, the company can leverage the low costs of the contract by signing the capacity contracts such that a commitment contract is already in place when demand during the drop is observed. We call this contracting strategy *Now-or-Never* and it is efficient when demand drops are expected.

**How Seasonal Modes Affect the Commitment Strategies.** Seasonal demand may manifest not as a surge or drop but rather with different seasonal modes, which is a typical setting for companies selling strongly seasonal products such as winter/summer sport equipment, holidays merchandise, etc. For simplicity of exposition, we assume that we have two modes: During the first half of the year, the distribution of the demand is expected to be low and during the second half, demand is expected to be larger.

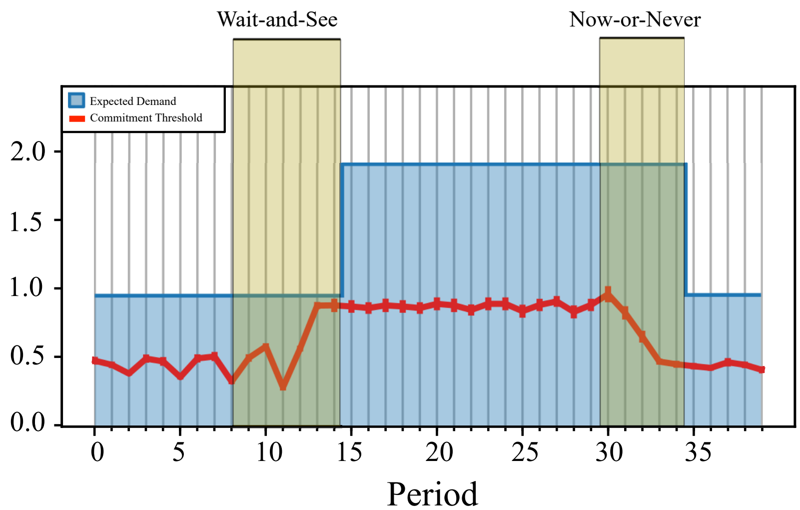


Figure 2.10. Threshold policy with seasonal modes with sharp change.

Similar to the results for seasonal demand surges and drops, the information is most valuable when contracts are coordinated with demand distribution changes. The demand visibility is leveraged if a capacity contract is signed when the surge period is observed for the first time. In other words, the company benefits from approaching the surge with no commitment in place. We numerically show this observation in Figure 2.10. For the periods prior to entering the surge (13-14), the threshold is high. This increases the likelihood of approaching the surge without a commitment in place. Hence, we observe that the *Wait-and-See* strategy is preferred when demand changes from low to high seasonal mode. On the other hand, the company desires to exit the surge with a commitment already in place. The company takes advantage of demand observation by signing a contract while demand is still high. We observe in periods (30-33) that the *Now-or-Never* strategy is preferred when the high demand mode ends. Observe that with seasonal modes, the dampening of oscillation prior to the first and the last period of the high mode occurs more slowly.

## 2.4. Conclusions and Further Research

Motivated by a collaboration with a major furniture company, we demonstrate how limited demand visibility can be used to efficiently select transportation procurement contracts in the presence of commitment. Our work is novel in its study of the combined effect of information and short commitment requirements in this environment. Particularly, we show that the availability of partial information fundamentally changes the policies compared to the case where only the demand distribution is known or the case where all the realizations of the demand are known in advance. Information regarding future demands can be used by the decision maker to determine whether or not a contract should be signed through a threshold policy. When a contract is signed, demand visibility allows the decision maker to choose customized contracts that adapt to the observations. In addition, when facing seasonal demand, the partial information is leveraged by properly coordinating the contracting epochs using tailored strategies. Applying these insights to the highest demand lanes of the furniture company, we find that considering demand information in contracting policies can reduce costs by 3-4% consistently across these lanes (see Appendix A.2).

The transportation procurement selection studied in this chapter describes a decision process faced in many business settings where companies have visibility of demand for the near future. For example, in addition to furniture enterprises, brick and mortar and e-commerce companies offering delivery services operate in a partial information setting as transportation needs are known within their delivery time window. Not only are these decisions quite common, they are also crucial to business success. According to the National Retail Federation, 60% of the online transaction included free shipping [**98**]. Hence, the competition in costs in this industry has resulted in a paradigm where the company typically assumes the shipping costs, increasing the cost of moving the products. For instance, as stated in CNN, Amazon's shipping costs reached \$1.5B in Q1 of 2016 while net profits reached \$513M [**58**]. Hence, consistently reducing 3-4% of shipping operations costs can have a significant impact on profitability. With the emergence of e-commerce and home delivery, the companies requiring highly efficient distribution of their products has increased in the past twenty years, which makes the study of the role of partial information and commitment grow in relevance.

This work opens multiple directions of future research. More general results can be found if one considers other forms of the costs $c$ and $c'$ (e.g. convex functions of capacity signed or commitment horizon, stochastic costs). From the transportation network perspective, we believe there is much work to be done in analysis of multiple lanes. Under that scenario, the capacity of the contract can be shared by different destinations. In the next chapter of this dissertation, Chapter 3, we study the impact of demand visibility on the capacity decisions in a network of transportation lanes, where multiple destinations can share transportation capacity (by means of multi-stop routes or flexible transportation contracts).

Another natural extension considers longer periods of demand visibility (i.e. $\gamma > 1$). In the transportation literature, expanded delivery windows (a form of expanding the information horizon) have been studied in terms of increased delivery density [**22**], and future research here could complement those studies. Wider delivery windows provide the company with more information, but it might come at a cost of decreased customer service. In Chapter 4, we explore the transactional data from the furniture company to study the impact of increasing delivery time windows on customers' behavior. A more detailed analysis of the value of additional information can be particularly useful in settings in which it is costly to gain additional demand visibility, or when visibility can be gained through increasing delivery windows.

CHAPTER 3

# The Value of Information and Flexibility with Temporal Commitments

*This chapter extends the results of Chapter 2 to a network with dedicated transportation contracts (single-leg lanes) and flexible contracts (two-leg lanes). In a network, demand visibility and flexible capacity are two hedging mechanisms against demand uncertainty for signing capacity contracts. With new technological innovations, temporal commitment of such contracts is shortening. Previous studies show that, with long temporal commitment, the two mechanisms are substitutes. We show that commitment imposed across few periods leads to new dynamics in which the two mechanisms act as both substitutes and complements.*

## 3.1. Introduction

In this chapter we study how partial demand visibility can help a firm improve its use of flexible capacity when the firm can choose to commit to capacity for a few periods. Our work is motivated by a furniture company that periodically signs contracts of relatively short temporal commitment duration (a few weeks) with transportation providers. In these contracts, the company commits to levels of dedicated and flexible transportation capacity that will be available for use for the duration of the commitment horizon.

While short term temporal commitments of few weeks are common in transportation where capacity (i.e. trucks) can be reassigned relatively easily, this has typically not been the case when sourcing traditional production capacity. In traditional procurement, temporal commitments are typically made for a year or longer because capacity cannot easily be adjusted. As a result, most sourcing literature has considered very long commitments. Our focus on relatively short commitment horizons is likely to become more relevant in the future as technological innovations make capacity easily redeployable (see [**45**] for contexts and literature). For example, Flexe, Stord, and Flow are companies that provide on-demand warehousing to their clients and allow commitments as short as a few weeks. In processing and manufacturing, there is a growing interest in mobile production capacity that is easily reconfigurable (e.g. [**14**] and [**88**]). Likewise, 3-D printing provides firms with production capacity that is easy to redeploy across clients. In these settings, the increased flexibility of the supply process (whether for warehousing or production) facilitates short term commitments. With shorter commitment horizon, demand visibility due to advance scheduling in services, make-to-order manufacturing (e.g. [**125**], [**78**], [**77**], [**15**]) and logistics (e.g. [**146**]) is likely to become more valuable.

There is an extensive literature that considers the use of flexible capacity when faced with demand uncertainty to improve performance (e.g. [**51**], [**128**] and [**60**]). There is also literature (see [**15**] and [**19**]) that shows how partial demand visibility can be leveraged to improve performance. This chapter combines these two streams and focuses on how partial demand visibility can allow a firm to suitably commit to and deploy flexible capacity given short term commitments.

To understand how demand visibility and flexible capacity may work together to address demand uncertainty, consider the middle mile of a transportation network connecting a hub to two spokes. A firm serves stochastic demand at two spoke locations by periodically distributing products from a central hub. The spoke cities can be served using dedicated truckload routes that only visit one spoke. When signing dedicated transportation contracts, the company must commit to a capacity level and to an origin-destination route, as shown in 3.1-(a). For flexible transportation contracts, the company commits to a capacity level, but has flexibility in the route destinations, including the more expensive option of multi-stop routes, as shown in 3.1-(a). The addition of multi-stop routes and the ability to change destinations based on realized demand leads to the higher cost of flexible contracts. If demand for the following week is visible, a flexible contract allows the company to load the flexible route based on observed demand at the spoke cities. Thus, the furniture company faces a trade-off between the higher cost of flexibility and its ability to use the flexible capacity as needed. Figure 3.1-(b) shows the decision sequence of the firm. At each decision point, the firm has *demand visibility* over the next period and must commit to contracts for a *commitment* duration of four periods. In each period, demand is first served using the committed capacity, with any unmet demand served using the spot market. The firm makes a new commitment decision after the four periods of commitment. Partial demand visibility along with short term commitment allows a firm to change its decision after each commitment horizon based on the new demand observed. As a result, short term commitment leads to findings that are different from the existing literature where commitment is either non-existent or very long.

Using a newsvendor network model we show that it is optimal to adapt (within bounds) the committed capacity for each commitment horizon based on the visible demand. Commitment to flexible capacity strongly depends on the visible demand at the two nodes. If visible demand at one node is high and visible demand at the other is low, it is optimal to commit to more flexible capacity and less dedicated capacity. Using demand visibility to adapt capacity commitments increases the average marginal value of flexible capacity by changing the commitment to flexible capacity based on visible demand.

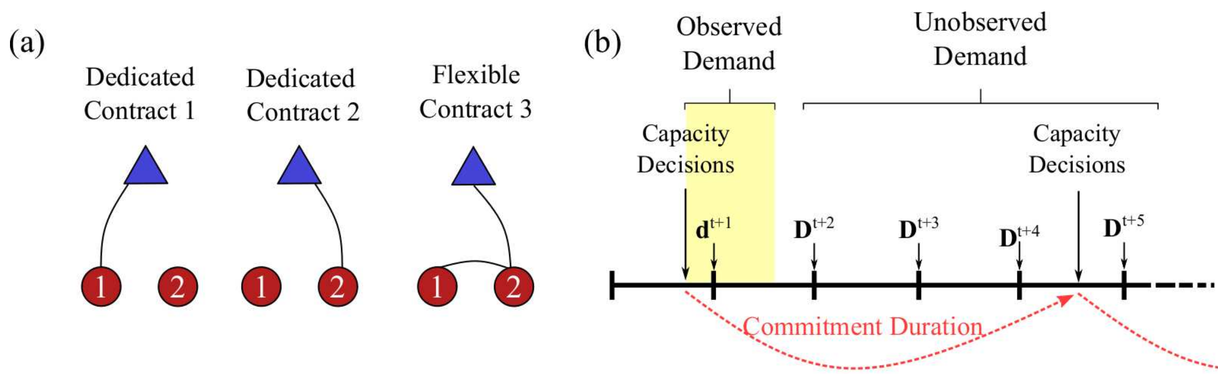


Figure 3.1. (a) Contracts available to supply nodes 1 and 2. (b) Time line of capacity decisions.

When a firm utilizes demand visibility and flexible capacity to hedge against uncertainty when committing to capacity, a natural question arises: Does the value of demand visibility increase or decrease in the presence of flexible capacity? In other words, are demand visibility and flexible capacity substitutes or complements? We find that demand visibility and flexible capacity can be either substitutes or complements depending on the cost of flexible capacity (see Figure 3.2). While demand visibility is always valuable, its value changes with the optimal commitment to flexible capacity. When the cost of flexible capacity is very high (far right white area of Figure 3.2), it is optimal to commit only to dedicated capacity. Clearly, demand visibility is valuable in this situation. At the other extreme, when the cost of flexible capacity is very low (far left striped area in Figure 3.2), it is optimal to commit to a high level of flexible capacity. The high level of flexible capacity reduces the value of demand visibility. In this situation, flexible capacity and demand visibility behave like substitutes, confirming the intuition found in previous studies. Demand visibility is most valuable between the two extremes (shaded area in Figure 3.2) where it is optimal to commit to both dedicated and flexible capacity. In this range of the cost of flexible capacity, flexible capacity and demand visibility behave like complements.

Our sensitivity analysis shows that the range over which partial demand visibility and flexible capacity are complements changes with the commitment horizon and the correlation of demand at the two spoke nodes. An increase in the commitment duration, or an increase in positive correlation of demand at the spokes, expands the range over which flexible capacity and demand visibility complement each other.

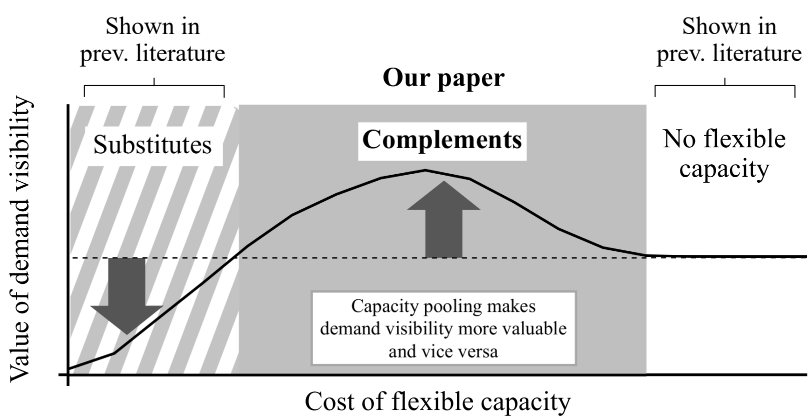


Figure 3.2. Value of demand visibility as a function the cost of flexible capacity.

Sections 3.2 and 3.3 introduce the basic model assumptions and related literature. Section 3.4 we describe our model and find basic insights regarding the optimal decisions with partial demand visibility and flexible capacity with commitment. We analyze and discuss the interplay between flexible capacity and partial demand information in Section

3.5. Sensitivity analysis with respect to key model parameters is presented in Section 3.6.

## 3.2. Context and Basic Model Assumptions

In this section we introduce the decision process faced by the company that motivates our study. We then discuss the assumptions used in our model.

### 3.2.1 Managerial Context

This chapter is motivated by a furniture company (as in [**19**]) with hundreds of showrooms across the United States that produces its furniture to order for delivery to customers' homes. At the point of purchase, customers choose a delivery date, typically, within one or two weeks. Once an order is scheduled, products are distributed from the hubs to the spokes, and from the spokes to the customers' home. Given that most orders are scheduled for delivery a week or more after purchase, the transportation group of the furniture manufacturer typically has visibility into scheduled delivery needs over the next week. Thus, the demand for transportation needs is visible one week into the future.

The company supplies most of its transportation needs with spot market options provided by companies such as XPO, UPS or U.S. Postal Service (USPS). The company also signs transportation procurement contracts of about four weeks at a lower cost than the spot market. By signing a longer term contract, the company commits to pay for a chosen weekly capacity level over the next four weeks while having demand visibility only a week into the future. The company faces the key decision of committing to capacity over the commitment horizon of four weeks.

In this situation, the firm has partial information in the form of demand visibility for a week, but must make commitments for four weeks. As a result, the firm has multiple commitment opportunities over the planning horizon of a quarter or a year. In relation to Chapter 2, we study the decision process of choosing committed dedicated and flexible capacity when the company decides to supply the transportation needs by committing to transportation contracts.

Figure 3.1-(a) depicts a region where the company serves two spokes (circles) from a single hub (triangle). When signing transportation contracts, the company must commit to a capacity level for each route. In Figure 3.1-(a), the spokes can be served by dedicated transportation capacity (a truck route that visits only one spoke node) and flexible transportation capacity (a truck route that visits both spoke nodes , either independently or with multi-stop routes). A dedicated capacity commitment can only be used to meet demand at a specific spoke node. In contrast, the flexible capacity commitment can be allocated in any amount up to the committed capacity to meet demand at either of the two spoke nodes visited. This ability to adjust route destinations and use multi-stop routes if necessary leads results in a higher cost for flexible capacity contracts. When making its capacity commitments, the company must trade off the higher cost of flexible capacity against the benefit of pooling offered by the flexible capacity.

Our work focuses on how flexible capacity and demand visibility, two mechanisms to hedge against demand uncertainty, interact in the presence of temporal commitment. In the following sections, we describe our basic assumptions and then discuss previous literature that explores the use of demand visibility (both in the presence and absence of commitment) and flexible capacity (prior literature has ignored commitment) to hedge against demand uncertainty.

### 3.2.2 The Basic Assumptions of the Model

Our notation is based on the nomenclature used in [**128**] and [**19**]. We use bold letters to denote vectors (e.g. $\mathbf{k}_i$). Element $j$ of $\mathbf{k}_i$ is $k_{i,j}$. $\mathbf{v} \succ (\prec)\mathbf{u}$ is element-wise comparison. Capital letters are reserved for random variables and lower case letters for their realizations. $\mathbb{E}_{\mathbf{D}}$ is the expectation with respect to $\mathbf{D}$. $\mathbb{R}^n$ is the $n$-dimensional real vector space and $\mathbb{R}^n_+$ is $\{\boldsymbol{v} \in \mathbb{R}^n | \boldsymbol{v} \succ 0\}$. $[x]^+$ is the positive part of $x$.

We consider a firm that uses a hub to supply the demand at two spokes, $n_1$ and $n_2$, as shown in Figure 3.1. In each period $t$, nodes $n_1$ and $n_2$ have demand realizations that we model as continuous random vectors $\boldsymbol{D}^t = (D_1^t, D_2^t)'$. We first assume that demands of different periods are independent and identically distributed (IID) and follow a known distribution $F(x, y)$. We allow the demand $D_1^t$ to be correlated with $D_2^t$, but not across time. Given that the value of information will increase with correlation across time, assuming independence across time is reasonable.

At the beginning of each commitment horizon, the firm decides the capacity level of three *network contracts* (see Figure 3.1-(a)). Dedicated-Contract-1 (contract 1): capacity commitment on truck traveling to node $n_1$. Dedicated-Contract-2 (contract 2): capacity commitment on truck traveling to node $n_2$. Flexible Contract (contract 3): capacity commitment on truck traveling to both nodes $n_1$ and $n_2$. The commitment decision is made for the *commitment horizon* $\tau$. The costs per unit per period of each contract are $c_1, c_2, c_3$, respectively. To avoid trivial solutions, we assume without loss of generality

$c_1 \leq c_2 < c_3 \leq c_1 + c_2$. Any portion of realized demand in a period that is in excess of the committed capacity is supplied using the spot market at a cost $c_s > c_3$.

As shown in Figure 3.1-(b), the decision maker observes the demand realizations for the next $\gamma$ periods (the *information horizon*) before making any commitment. In the *null information* scenario ($\gamma = 0$), no demand is visible. In the *partial demand information* scenario ($\gamma \geq 1$), demand is visible for the next $\gamma$ periods. In this case, we say that partial demand information is of the form of partial demand visibility. With this information, the decision maker commits to capacity for each contract for a duration of $\tau$ periods (commitment horizon).

Our stylized model has three assumptions that simplify the analysis and exposition. Each assumption, however, can be relaxed while maintaining the core insights. Our first assumption is to focus on a simple network with two demand nodes where the spot market cost of serving each node is the same. Our results (Theorems 2-8) can easily be translated to the case when spot market costs for the two nodes are different. Our insights also extend to networks with more than two nodes that share flexible capacity. Second, we assume that the demand in each period has known, stationary and identical distribution. This assumption allows us to easily compare the solution of the partial information case with the null information case. In Section 3.7, we generalize this assumption and extend our results to the case where demand is not stationary and the case where distribution is unknown. Our third assumption is to fix the information horizon to a single period ($\gamma = 1$). The intuition developed when the information horizon is a single period holds when this assumption is relaxed.

We support our analytical findings with numerical experiments where the model parameters are calibrated with realistic values based on data from the furniture company. The ratio between per unit dedicated capacity costs and spot market costs is $c_i/c_s \approx 0.3$ (the ratio of $c_i/c_s$ ranged between 0.2 and 0.4 for the company). We use a coefficient of variation of demand of 0.6 which is within the range 0.5-0.7 for the company. Our commitment horizon $\tau = 4$ weeks and information horizon $\gamma = 1$ week are close to the scenario faced by the company. Demand between nearby nodes also showed positive demand correlation ranging from $\rho = 0.1$ to 0.3. We explore the impact of key model parameters on our results in Section 3.6.

### 3.2.3 Base Case: The Null Information Model With Flexible Capacity ($\gamma = 0$)

In this section we set up the null information model, i.e., the case without demand visibility ($\gamma = 0$). We build on the model given in [**128**] to obtain optimal capacities for dedicated and flexible contracts when the distribution of demand is known but there is no demand visibility. Since demand is stationary, all periods are identical and we can focus on the case $\tau = 1$.

We model the capacity decisions with a set of *ex-ante* decisions at the start of each commitment horizon and a set of *ex-post* decisions for each period after demand is observed. The *ex-ante* decision minimizes total expected cost by identifying the desired committed capacity $k_1, k_2$ and $k_3$ for each of the three contracts at unit costs per period $c_1, c_2$ and $c_3$, respectively. Once demand is realized in each period, the *ex-post* decision optimally uses the committed capacities and the spot market to supply the demand at both nodes. We denote by $y_1$ ($y_2$) the amount demand at node 1 (2) supplied using the dedicated contract 1 (2). We denote by $y_3$ ($y_4$) the quantity supplied to node 1 (2) using the flexible contract. These quantities can be interpreted as the units flowing through each edge of Figure 3.3-(a). The quantities shipped using the spot market are obtained as $[d_1 - (y_1 + y_3)]$ and $[d_2 - (y_2 + y_4)]$ for nodes $n_1$ and $n_2$ respectively. We denote $\mathbf{y} = (y_1, y_2, y_3, y_4)' \in \mathbb{R}^4_+$.

For a committed capacity vector $\mathbf{k} = (k_1, k_2, k_3)' \in \mathbb{R}^3_+$ and demand realization $\mathbf{d} = (d_1, d_2)' \in \mathbb{R}^2_+$, the *ex-post* problem is to find a flow vector $\mathbf{y}$ that minimizes the spot market cost to meet demand (because the cost of committed capacity is a sunk cost for each period). The *ex-post* problem of minimizing the spot market costs given demand realizations $(d_1, d_2)$ is formulated as follows:

$$\pi(\mathbf{k}, \mathbf{d}) = \min_{\mathbf{y} \in \mathbb{R}^4_+} \quad c_s(d_1 - y_1 - y_3) + c_s(d_2 - y_2 - y_4) \tag{3.1}$$

subject to:

$$\begin{aligned} y_1 + y_3 &\le d_1 \\ y_2 + y_4 &\le d_2 \end{aligned} \tag{3.2}$$

$$\begin{aligned} y_1 &\le k_1 \\ y_2 &\le k_2 \\ y_3 + y_4 &\le k_3. \end{aligned} \tag{3.3}$$

The optimal solution of $\pi(\mathbf{k}, \mathbf{d})$, the ex-post problem, partitions the sample space of demand $\boldsymbol{d} \in \mathbb{R}^2$ into four disjoint regions such that the dual variables are constant in each region and the solution of $\pi(\mathbf{k}, \mathbf{d})$ is fully characterized. These regions are defined below and shown in Figure 3.3-(b).

**Definition 2.** *We define the set of regions $\Omega_i(\boldsymbol{k})$ $(i = 0, I, II, III)$ as a function of the vector of capacity vector $\boldsymbol{k}$ as follows:*

- $\Omega_0(\boldsymbol{k}) = \{\boldsymbol{d} \in \mathbb{R}^n_+ | d_1 + d_2 \leq k_1 + k_2 + k_3, d_1 \leq k_1 + k_3, d_2 \leq k_2 + k_3\}$
- $\Omega_I(\boldsymbol{k}) = \{\boldsymbol{d} \in \mathbb{R}^n_+ | d_1 \leq k_1, d_2 \geq k_2 + k_3\}$
- $\Omega_{II}(\boldsymbol{k}) = \{\boldsymbol{d} \in \mathbb{R}^n_+ | d_2 \leq k_2, d_1 \geq k_1 + k_3\}$
- $\Omega_{III}(\boldsymbol{k}) = \{\boldsymbol{d} \in \mathbb{R}^n_+ | d_1 + d_2 \geq k_1 + k_2 + k_3, d_1 \geq k_1 + k_3, d_2 \geq k_2 + k_3\}$.

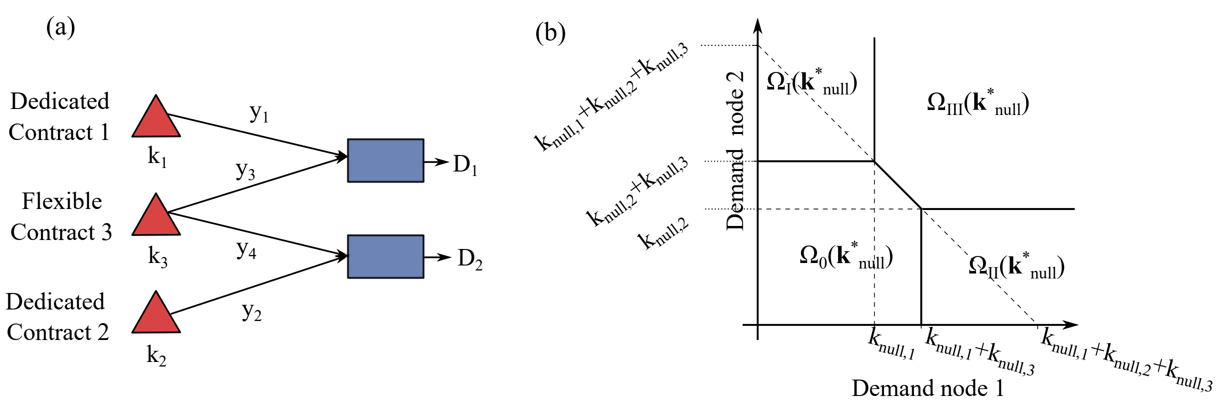


Figure 3.3. (a) Basic contracts in a newsvendor network. (b) Adaptation of Figure 1 of [**128**]: The four regions where the solution of $\pi(\mathbf{k}, \mathbf{d})$ is fully characterized and the shadow prices are constant.

Consider the *ex-ante* problem of determining the contract capacities $k_1, k_2$ and $k_3$. Let $\mathbf{D} = (D_1, D_2)'$ be the random vector of demands and $\mathbf{c} = (c_1, c_2, c_3)'$ the vector of contract costs. The optimal committed capacity minimizes the cost of committing to the contracts and the expected cost incurred from the spot market in Eq. (3.4).

$$V(\mathbf{k}) = \mathbb{E}_{\mathbf{D}}\, \pi(\mathbf{k}, \mathbf{D}) + \mathbf{c}'\mathbf{k} \tag{3.4}$$

We refer to the minimization of Eq. (3.4) as the *null information problem* and denote its solution by $\mathbf{k}^*_{null}$. While there is no closed form solution to $\mathbf{k}^*_{null}$, Proposition 1 in [**128**] provides First Order Conditions (FOC), which are necessary and sufficient for optimality because the minimization problem defined in Eq. (3.4) is convex and differentiable. The rest of the chapter studies how partial demand visibility changes capacity commitments relative to the null information model.

## 3.3. Related Literature

In this section we review literature related to flexible capacity and partial demand visibility.

### 3.3.1 Capacity Pooling in the Literature

Procurement and capacity management within a network under null information (no demand visibility) has been studied by several authors including [**51**], [**76**], [**62**] and [**69**]. [**51**] provides necessary and sufficient conditions for investments in flexible capacity to be optimal. Our work focuses on the decision of selecting capacity contracts in a network when demand is uncertain over a commitment horizon, and we model the procurement decision as a newsvendor network. We adapt the results in [**128**] to describe the null information scenario and focus on understanding how partial demand visibility changes the structure of the optimal policy. [**130**] and [**129**] study the newsvendor network model under dynamic environments and risk considerations. The value of flexible capacity in conjunction with other mechanisms or business models is also studied in the literature. For instance, [**32**] studies the impact of pricing decisions on the value of flexibility with responsive demand. [**121**] shows that the value of flexible capacity is reduced when the decision maker is risk averse and suppliers are not reliable. In multimarket network design, [**85**] describes how demand uncertainty increases the value of centralized procurement strategies due to flexible capacity. The authors also show that demand volatility and price differentiation change the optimal configuration of the centralized network. [**59**] and [**46**] show how incentives for choosing flexibility change with demand competition in a single and a multi market duopoly, respectively. The studies of [**33**] and [**34**] consider the interaction of flexible capacity and financial hedging. The former shows that flexible contracts and financial hedging are complements when demand is positively correlated and are substitutes when the correlation is negative. The latter considers how willingness to invest in flexibility changes when financial debt to raise capital is incorporated into the model.

Our major contribution to this literature is to show how the availability of partial information in the presence of commitment impacts the optimal utilization of flexible capacity.

### 3.3.2 Demand Visibility in the Literature

Many papers have considered the scenario where the decision maker knows the demand distribution but has no demand visibility when capacity decisions are made at the beginning of the planning horizon (e.g. [**6**] and [**93**]). [**19**] considers the impact of partial demand visibility when the commitment horizon is not too long and shows that partial information is an effective hedge against demand uncertainty in the presence of commitment.

Several papers have studied the value of partial information in supply chains (see [**132**] for a review). In transportation systems, [**145**] and [**146**] numerically show that demand information can increase profits for pick up and delivery with trucks by improving schedules and route planning. [**147**] studies the value of information in maritime container operations and [**106**] studies the effect of advance demand information and pricing in a newsvendor model. [**78**] considers the value of knowing orders in advance in a single-stage make-to-stock queue. [**135**] develops a dynamic stochastic inventory model with fulfilment date flexibility, and numerically quantifies the cost reduction when demand information is used to alleviate the impact of demand uncertainty. [**49**] and [**50**] study how availability of information changes the interaction between a seller and a buyer when demand forecasts updates are allowed. Among others, [**5**], [**103**] and [**123**] study the value of information when it is used to improve demand forecasts.

Most of the literature focuses on a single location without a commitment horizon. In contrast, our work considers a network with two locations (that can also be interpreted as products) and focuses on how demand visibility impacts the optimal use of flexible capacity in the presence of commitment.

### 3.3.3 Combining Pooling and Demand Visibility Without Commitment

We highlight three studies that combine demand information and flexible capacity in the absence of commitment. [**100**] considers Advanced Demand Information (ADI) in a system with multiple retailers supplied by a central warehouse. The author finds structural results of a lower bound relaxation and numerically shows that the availability of demand information increases the value of postponement (by moving the central warehouse closer to the retailers). In a similar vein we show that demand visibility can increase the value of flexible capacity in the presence of commitment. [**77**] studies the value of ADI for $N$ independent locations versus $N$ locations supplied by a single flexible resource. According to the authors: "*ADI is more valuable when there is a lot of demand variability that can be removed using ADI. Resource sharing seems to make ADI relatively less valuable precisely for this reason*". In contrast, in the presence of commitment, we show that the availability of shared flexible capacity makes demand visibility more valuable. [**15**] studies the impact of ADI in a multiproduct system. The authors show that when the flexible capacity is shared by all products, information about the total demand increases the decision maker's willingness to invest in flexible capacity but information about individual product shares decreases the decision maker's willingness to invest in flexible capacity. In contrast, we show that in the presence of commitment, knowledge of individual product demand can increase the decision maker's willingness to invest in flexible capacity.

Our findings match with some previous studies where visibility and flexibility are shown to behave like complements in other contexts. [**31**] studies how flexible resources help decision makers improve the estimation of demand distributions by reducing the probability of demand censoring occurrences. Hence, flexible contracts complement the learning capabilities. We show that even without censoring demand, flexible contracts can still complement demand visibility in the presence of commitment. [**75**] studies sourcing from an unreliable supplier with a random lead time and a reliable but more expensive supplier with a deterministic and shorter lead time. The author numerically finds that lead time visibility of the unreliable source adds value by allowing better use of the reliable source. The empirical work in [**114**] shows that organizational flexibility increases the value obtained from analytic capability coupled with visibility. In the context of procurement with commitment, we show demand visibility adds value by allowing better use of flexible capacity.

## 3.4. The Partial Demand Information Model in a Network

In this section we model the interaction between flexible capacity and partial demand information under commitment. We first formulate the general model where both dedicated and flexible capacity are available along with partial demand information under commitment. We then analyze a special case where only dedicated capacity is used. We then extend the analysis to the general case where both dedicated and flexible capacity are available.

### 3.4.1 The Cost Function Under Partial Demand Information

Recall that the base model of Section 3.2.3 considers the problem where there is no demand visibility (the null case) but both dedicated and flexible capacity are available under commitment. We build on the base model by including partial demand visibility. We consider the decision sequence outlined in Figure 3.1-(b). For information horizon $\gamma = 1$, the decision maker observes the next period demand $\boldsymbol{d} = (d_1, d_2)'$ at the two nodes and then commits to a capacity vector $\boldsymbol{k} = (k_1, k_2, k_3)'$ for the commitment horizon. The expected cost incurred over the commitment horizon consists of two parts:

(1) *Visible cost.* The cost over the period with visible demand: $\text{Cost}_{visible}(\mathbf{k}, \mathbf{d}) = \pi(\mathbf{k}, \mathbf{d}) + \mathbf{c}'\mathbf{k}$.

(2) *Expected cost.* The expected cost incurred over the remaining $\tau - 1$ periods in the commitment horizon where $V(\mathbf{k})$ of Equation (3.4) is the expected per-period cost: $\text{Cost}_{expected}(\mathbf{k}) = (\tau - 1)V(\mathbf{k})$.

Therefore, the *total expected cost* over the commitment horizon $\tau$ is obtained as follows:

$$\text{Cost}_{total}(\mathbf{k}, \mathbf{d}) = \text{Cost}_{visible}(\mathbf{k}, \mathbf{d}) + \text{Cost}_{expected}(\mathbf{k}) = \pi(\mathbf{k}, \mathbf{d}) + \mathbf{c}'\mathbf{k} + (\tau - 1)V(\mathbf{k}). \tag{3.5}$$

The total expected cost function in Equation (3.5) is non-smooth, continuous and convex with respect to $\mathbf{k}$. Let $\mathbf{k}^*(\mathbf{d})$ be the capacity vector that minimizes total expected cost.

Observe that the term $(\tau - 1)V(\mathbf{k})$ dominates in (3.5) when the commitment duration $\tau$ is very long. In this case, the optimal solution is very close to that of the null information scenario, as in Section 3.2.3. In the limit, as $\tau \to \infty$, $\mathbf{k}^*(\mathbf{d}) \to \mathbf{k}^*_{null}$. Thus, the base case without demand visibility is a limiting case of our model as the commitment horizon $\tau$ becomes longer.

When commitment duration is very short ($\tau \to 1$), the problem is similar to a deterministic single period decision with the visible cost being the dominant component of total expected cost. In this case, the decision maker should behave myopically by not investing in flexible capacity (because it is more expensive) and setting the committed dedicated capacity exactly equal to the demand that is observed at each node.

When commitment duration is between these two extremes ($1 < \tau < \infty$), the decision maker faces the trade off between myopically adapting to the demand observed, while still considering expected cost in the remaining $\tau - 1$ periods where demand is not visible. We first study the special case where the network has only dedicated contracts (no flexible capacity), and provide a closed form solution. We then consider the general case where the decision maker can use a combination of dedicated and flexible capacity.

### 3.4.2 A Special Case: A Network with Dedicated Contracts

In this section we analyze the case where each node can only be served using dedicated capacity. This allows each node $n_i$ to be considered independently.

Assume that demand in each period is IID with known marginal distribution $F(x)$. Under null information, the optimal committed dedicated capacity for node $n_i$ (denoted by $k^*_{null,i}$) is given by the newsvendor fractile: $k^*_{null,i} = F^{-1}(\frac{c_s - c_i}{c_s})$. Under null information, it is optimal to commit to the same capacity level at the beginning of each commitment horizon.

Under partial demand information, [**19**] shows that optimal committed capacity adapts (within bounds) to the visible demand $(d_1, d_2)$. For example, if visible demand at node $i$, $d_i$ is quite low, the optimal committed dedicated capacity is smaller than the capacity under null information ($k^*_{null,i}$). Proposition 7 formalizes this result to the case of a dedicated network in the absence of flexible capacity.

**Proposition 7.** *In the absence of flexible capacity, the optimal dedicated capacity to supply node $n_i$ ($i = 1, 2$) is determined by the observed demand $d_i$ as follows: Define $\psi_i^l := F^{-1}\left(\frac{c_s - (c_i + \frac{c_i}{(\tau-1)})}{c_s}\right)$, and $\psi_i^u := F^{-1}\left(\frac{c_s - (c_i - \frac{c_s - c_i}{(\tau-1)})}{c_s}\right)$. Then,*

$$k_i^*(d_i) = \begin{cases} \psi_i^l & \textit{if } d_i < \psi_i^l \\ d_i & \textit{if } \psi_i^u \geq d_i \geq \psi_i^l \\ \psi_i^u & \textit{if } d_i > \psi_i^u. \end{cases} \tag{3.6}$$

When visible demand is between the lower and upper thresholds ($\psi_i^l$ and $\psi_i^u$ respectively), the optimal committed dedicated capacity ($k_i^*(d_i)$) equals observed demand. When the observed demand is higher (lower) than the threshold $\psi_i^u$ ($\psi_i^l$), the committed capacity equals $\psi_i^u$ ($\psi_i^l$). In other words, the optimal committed capacity adapts to the visible demand within bounds. Observe from Equation (3.6) that as the duration of the commitment horizon $\tau$ increases, $\psi_i^l$ and $\psi_i^u$ converge to $k^*_{null,i} = F^{-1}(\frac{c_s - c_i}{c_s})$, reducing the decision maker's willingness to adapt to observed demand.

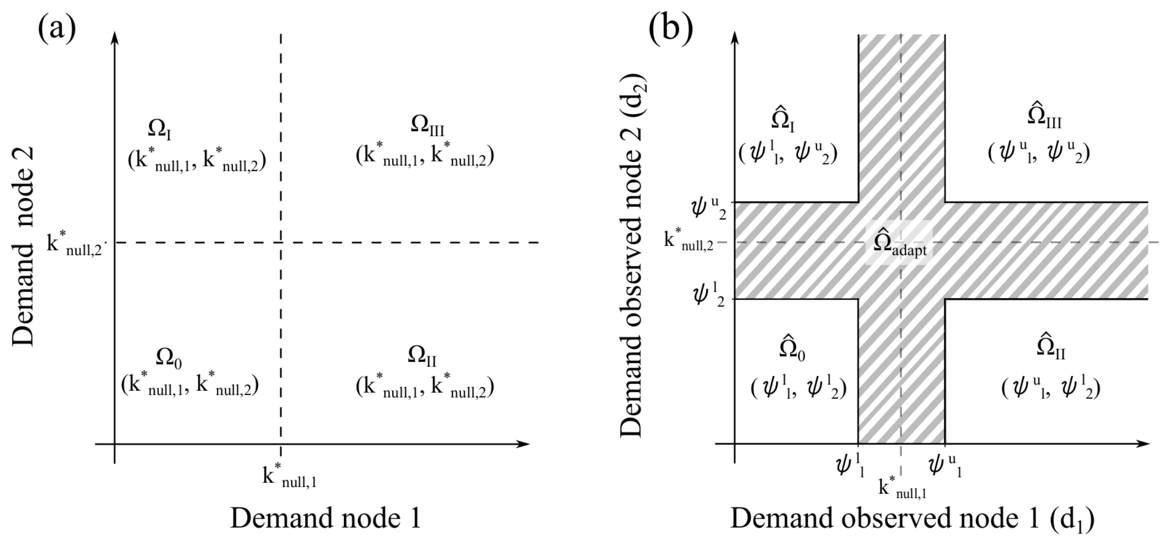


Figure 3.4. (a) Sample space with null information and no flexible capacity. (b) Sample space of observed demand with partial information and no flexible capacity. The striped region shows $\hat{\Omega}_{adapt}$.

In Figure 3.4 we plot the optimal dedicated capacity for both the null and partial information cases. Figure 3.4-(a) shows the optimal policy for the null information case where the sample space is divided into four regions, $\Omega_0, \Omega_I, \Omega_{II}$ and $\Omega_{III}$ based on whether the realized demand is higher or lower than the committed dedicated capacity $(k^*_{null,1}, k^*_{null,2})$.

Figure 3.4-(b) shows the optimal policy with partial demand visibility. In this case, the sample space of observed demand is divided into five regions, $\hat{\Omega}_{adapt}, \hat{\Omega}_0, \hat{\Omega}_I, \hat{\Omega}_{II}$ and $\hat{\Omega}_{III}$, that determine the capacity committed for each node depending upon visible demand. We show in parenthesis the optimal committed capacity vector $(k^*_1(d_1), k^*_2(d_2))$ in each region as determined by Proposition 7. When observed demand is within region $\hat{\Omega}_{adapt}$ (the striped region in Figure 7-b), the optimal committed capacity equals visible demand: $k^*_i(d_i) = d_i$, i.e., the optimal committed capacity changes with observed demand in region $\hat{\Omega}_{adapt}$.

If observed demand is outside $\hat{\Omega}_{adapt}$, the decision maker bounds the adaptation by choosing the optimal committed capacity of each contract to be the lower threshold $\psi^l_i$ for $i = 1, 2$ if demand is below $\psi^l_i$, or to be the upper threshold $\psi^u_i$ for $i = 1, 2$ if demand is above $\psi^u_i$. The region with demand below $\psi^l_i$ (above $\psi^u_i$) for both nodes is denoted by $\hat{\Omega}_0$ ($\hat{\Omega}_{III}$). If the demand observed at node 1 (2) is above $\psi^u_1$ ($\psi^u_2$) and at node 2 (1) is below $\psi^l_2$ ($\psi^l_1$), committed capacity at node 1 (2) is $\psi^l_1$ ($\psi^l_2$) and at node 2 (1) is $\psi^u_2$ ($\psi^u_1$). This corresponds to region $\hat{\Omega}_I$ ($\hat{\Omega}_{II}$). Note that the regions $\hat{\Omega}_0, ..., \hat{\Omega}_{III}$ define the boundary of $\hat{\Omega}_{adapt}$. Each region limits the extent to which the decision maker adapts optimal committed capacities to observed demand. As the commitment horizon $\tau$ increases, $\hat{\Omega}_{adapt}$ shrinks and the regions in Figure 3.4-(b) approach the regions in Figure 3.4-(a). In the limit, as $\tau$ approaches to infinity, they are identical.

In a dedicated network without flexible capacity, the decisions for the two nodes are independent and observed demand of one node does not impact the dedicated capacity commitment at the other node. In contrast, as we will show in the following section, in a network with both dedicated and flexible capacity, not only does the optimal flexible commitment depend on $d_1$ and $d_2$ but also the optimal dedicated commitment of each node depends on the demand observed at both nodes.

### 3.4.3 The Partial Demand Information Model with Dedicated and Flexible Commitments

Our goal is to understand how optimal capacity decisions with partial demand visibility deviate from the null information solution in the presence of dedicated and flexible capacity. In what follows, we assume that $\mathbf{c} \succ \frac{c_s}{\tau}$. (In Appendix B.3, we analyze the cases where this assumption does not hold.) We start with two technical definitions:

**Definition 3** (**Boundary Capacities**). *Define the boundary capacities* $\boldsymbol{\psi}^i$ *for* $i = \{0, I, II, III\}$ *as follows:*

$$\boldsymbol{\psi}^i = \{\mathbf{k} \in \mathbb{R}^3_+ \,|\exists \boldsymbol{\nu} \in \mathbb{R}^3_+ \; s.t. \; 0 = \mathbf{c} - \boldsymbol{\nu} - c^s\mathbf{v^i} + (\tau - 1)\nabla_{\tilde{\mathbf{k}}=\mathbf{k}} V(\tilde{\mathbf{k}}) \; and \; \boldsymbol{\nu}'\mathbf{k} = 0\} \tag{3.7}$$

*where* $\mathbf{v^0} = (0,0,0)'$, $\mathbf{v^I} = (0,1,1)'$, $\mathbf{v^{II}} = (1,0,1)'$ *and* $\mathbf{v^{III}} = (1,1,1)'$.

**Definition 4** (**Bounding and Adaptation Regions**). *For* $i = \{0, I, II, III\}$, *define the bounding regions as* $\hat{\Omega}_i = \Omega_i(\boldsymbol{\psi}^i)$, *where* $\Omega_i(\boldsymbol{k})$ *is as in Definition 2. Define the total bounding region as* $\hat{\Omega}_{bound} = \bigcup_{i=\{0,I,II,III\}} \hat{\Omega}_i$; *and the adaptation region as* $\hat{\Omega}_{adapt} = \mathbb{R}^2_+ / \hat{\Omega}_{bound}$.

Theorem 2 shows that there always exists an adaptation region $\hat{\Omega}_{adapt}$ with a boundary defined by $\hat{\Omega}_0, \hat{\Omega}_I, \hat{\Omega}_{II}$, and $\hat{\Omega}_{III}$, as shown in Figure 3.5. Theorem 2 generalizes Proposition 7 to include flexible capacity and also generalizes the results of [**128**] to include partial demand visibility. The proof is provided in Appendix B.1.

**Theorem 2.** *[**Characterizing the Adaptation Region**] Assume the availability of both dedicated and flexible capacity, demand visibility over the next period, and a commitment horizon of $\tau$. Let $\mathbf{c} \succ \frac{c_s}{\tau}$. Then, for $i = \{0, I, II, III\}$:*

*(1) $\hat{\Omega}_i \subset \Omega_i(\mathbf{k}^*_{null})$. Thus, there exists a non-empty adaptation region, $\hat{\Omega}_{adapt}$.*

*(2) The optimal capacity vector $\mathbf{k}^*(\mathbf{d}) = \boldsymbol{\psi}^i$ is constant for all points $\mathbf{d} \in \hat{\Omega}_i$. Capacity decisions are constant for all demand observations within each bounding region $\hat{\Omega}_i$.*

Figure 3.5 shows the partition of demand space $\boldsymbol{D}$ into the regions $\hat{\Omega}_0, \hat{\Omega}_I, \hat{\Omega}_{II}, \hat{\Omega}_{III}$ and $\hat{\Omega}_{adapt}$ according to Theorem 2. Observe that Figure 3.5 generalizes Figure 3.3 to include partial demand visibility and Figure 3.4-(b) to include flexible capacity.

As a result of partial demand visibility, the first part of Theorem 2 guarantees the existence of $\hat{\Omega}_{adapt}$ (striped region in Figure 3.5) where optimal committed capacity varies with observed demand. The adaptation is bounded by the optimal committed capacity in the bounding regions $\hat{\Omega}_0, \hat{\Omega}_I, \hat{\Omega}_{II}, \hat{\Omega}_{III}$ ($\hat{\Omega}_{bound}$), which are strict subsets of the regions $\Omega_i(\mathbf{k}^*_{null})$ (defined in Figure 3.3) for the null case. The second part of Theorem 2 states that for $i \in \{0, I, II, III\}$, the optimal committed capacity is constant at $\boldsymbol{\psi}^i$ as long as observed demand $(d_1, d_2)$ is within the bounding region $\hat{\Omega}_i$.

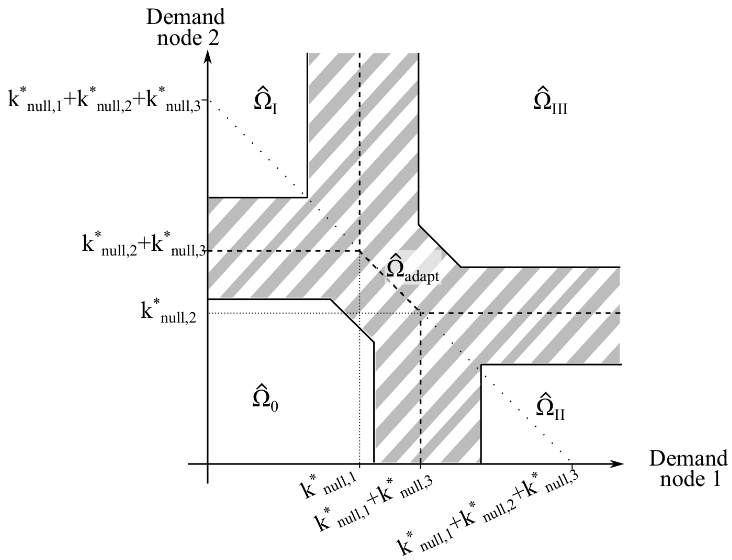


Figure 3.5. Sample space of demand partitioned into $\hat{\Omega}_0, \hat{\Omega}_I, \hat{\Omega}_{II}, \hat{\Omega}_{III}$ and $\hat{\Omega}_{adapt}$.

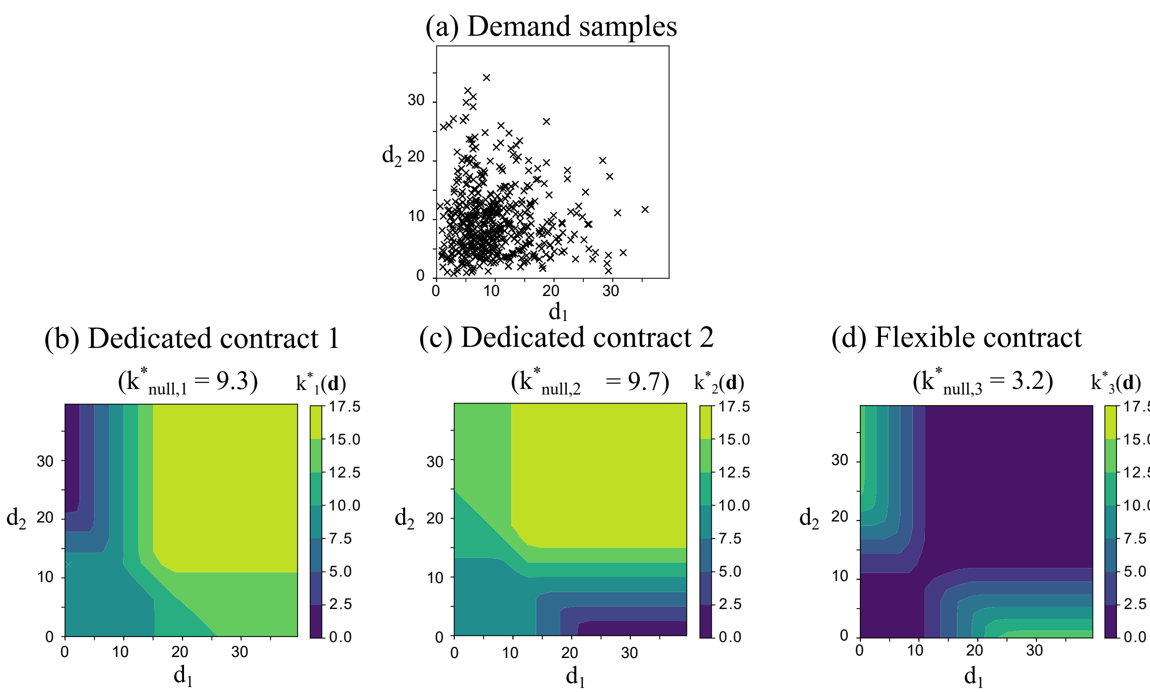


Figure 3.6. Numerical example of optimal capacity vector. (a) Demand samples. (b) The optimal capacity of the dedicated contract 1. (c) The dedicated contract 2. (d) The flexible contract.

Figure 3.6 numerically shows the change in capacity commitments with observed demand. Figure 3.6-(a) shows 250 randomly sampled demand vectors from binomial gamma distributions with means $\boldsymbol{\mu} = (10, 10)'$ and $\boldsymbol{\Sigma} = \begin{pmatrix} 6^2, 0 \\ 0, 6^2 \end{pmatrix}$. In this experiment, we use capacity costs $\boldsymbol{c} = (0.3, 0.3, 0.45)'$, spot market cost $c_s = 1$ and commitment horizon $\tau = 4$. Figures 3.6-(b),(c) and (d) respectively illustrate how the optimal committed capacities $k_1^*(\boldsymbol{d})$, $k_2^*(\boldsymbol{d})$ and $k_3^*(\boldsymbol{d})$, change with observed demand. Our computational experiments first replicate a version of Figure 3.5 to ensure the existence of all regions and then show how optimal commitments to dedicated and flexible capacity change with observed demand.

In Figure 3.6-(b), (c) and (d), observe the existence of region $\hat{\Omega}_{III}$ for $d_1$ and $d_2$ high enough (roughly $d_1 \geq 20$ and $d_2 \geq 22$) where the optimal capacity commitment remains constant at $\boldsymbol{\psi}^{III} = (15.1, 13.3, 0.0)'$. Similarly, observe the existence of region $\hat{\Omega}_0$ for low values of $d_1$ and $d_2$ (roughly $d_1 \leq 10$ and $d_2 \leq 10$) where the optimal capacity commitment stays constant at $\boldsymbol{\psi}^0 = (9.2, 9.2, 1.7)'$. We also observe the existence of region $\hat{\Omega}_I$ for low values of $d_1$ and high values of $d_2$ (roughly $d_1 \leq 2$ and $d_2 \geq 25$) where the optimal capacity commitment stays constant at $\boldsymbol{\psi}^I = (0.0, 10.9, 12.3)'$. We also observe the existence of region $\hat{\Omega}_{II}$ for high values of $d_1$ and low values of $d_2$ (roughly $d_1 \geq 25$ and $d_2 \leq 2$) where the optimal capacity commitment stays constant at $\boldsymbol{\psi}^{II} = (12.5, 0.0, 11.8)'$.

The bounding regions define the adaptation region within which the optimal committed capacity adapts to observed demand. For example, the optimal committed capacity when $d_1 = 22$ and $d_2 = 5$ is very different from the optimal committed capacity when $d_1 = 22$ and $d_2 = 22$. In Theorem 3 we show that the optimal committed capacity in the adaptation region is always in the convex hull of the optimal capacities in the bounding regions.

**Theorem 3.** *[**The Optimal policy**] Assume the conditions of Theorem 2. Let* $\mathbf{d} \in \mathbb{R}^2_+$ *be the vector of the observed next period demand. Then for* $i \in \{0, I, II, III\}$*:*

- *If* $\mathbf{d} \in \hat{\Omega}_i \subset \Omega_{bound}$*, the optimal committed capacity is* $\boldsymbol{\psi}^i$ *(*$\mathbf{k}^*(\mathbf{d}) = \boldsymbol{\psi}^i$*).*
- *If* $\mathbf{d} \in \Omega_{adapt}$*, the optimal committed capacity* $\mathbf{k}^*(\mathbf{d})$ *can be obtained as a convex combination of the boundary capacities, i.e.,*

$$\mathbf{k}^*(\mathbf{d}) = \sum_{i \in \{0,I,II,III\}} \alpha_i(\mathbf{d}) \boldsymbol{\psi}^i$$

*with:*

$$\alpha_i(\mathbf{d}) \geq 0, \quad {\scriptstyle \forall i \in \{0,I,II,III\}} \quad \textit{and} \quad \sum_{i \in \{0,I,II,III\}} \alpha_i(\mathbf{d}) = 1.$$

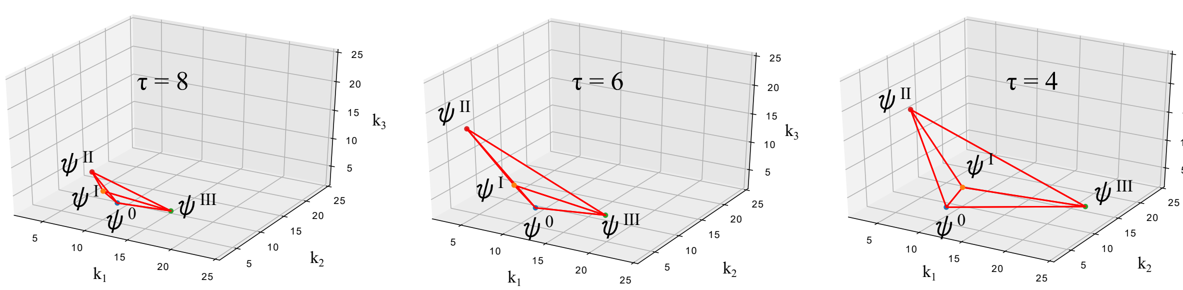


Figure 3.7. Convex hull delimited by the boundary capacities.

Theorem 3 generalizes the result in Proposition 7 to include flexible capacity. Figure 3.7 shows the convex hull generated by the boundary capacities for different commitment horizons $\tau$ for the numerical experiment described in Figure 3.6. Observe that it is optimal to expand the adaptation of committed capacities to observed demand as the commitment horizon $\tau$ shortens because partial demand visibility for a period is more valuable over shorter commitment horizons.

Proposition 8 describes the change in optimal committed capacities in the bounding regions relative to the null information solution $\boldsymbol{k}^*_{null}$ when partial demand visibility is included.

**Proposition 8.** *Let $\psi^i_j$ be the element $j$ of the boundary capacity $\boldsymbol{\psi}^i$ for $i \in \{0, I, II, III\}$ (Definition 2) and $k^*_{null;j}$ the $j$ component of the null information solution $\boldsymbol{k}^*_{null}$. If $\mathbf{c} \succ \frac{c_s}{\tau}$ the following inequalities hold:*

- *If observed demands at both nodes are low or both are high:*
  - *(1) In region $\hat{\Omega}_0$ (where both $d_1$ and $d_2$ are low) with partial demand visibility, it is optimal to commit to less total capacity (dedicated plus flexible capacity) relative to $\boldsymbol{k}^*_{null}$. This also holds for total capacity that can be used for each node.*

$$\psi^0_1 + \psi^0_2 + \psi^0_3 < k^*_{null;1} + k^*_{null;2} + k^*_{null;3} \ \ \textbf{\textit{and}}$$
$$\psi^0_1 + \psi^0_3 < k^*_{null;1} + k^*_{null;3} \ \ \textbf{\textit{and}} \ \ \psi^0_2 + \psi^0_3 < k^*_{null;2} + k^*_{null;3}.$$

  - *(2) In region $\hat{\Omega}_{III}$ (where both $d_1$ and $d_2$ are high) with partial demand visibility, it is optimal to commit to more total capacity (dedicated plus flexible capacity) relative to $\boldsymbol{k}^*_{null}$. This also holds for total capacity that can be used for each node.*

$$\psi^{III}_1 + \psi^{III}_2 + \psi^{III}_3 > k^*_{null;1} + k^*_{null;2} + k^*_{null;3} \ \ \textbf{\textit{and}}$$
$$\psi^{III}_1 + \psi^{III}_3 > k^*_{null;1} + k^*_{null;3} \ \ \textbf{\textit{and}} \ \ \psi^{III}_2 + \psi^{III}_3 > k^*_{null;2} + k^*_{null;3}.$$

- *If observed demand at one node is low while at the other is high:*

  *(3) In region $\hat{\Omega}_I$ (where $d_1$ is low and $d_2$ is high) with partial demand visibility, it is optimal to commit to less dedicated capacity for node 1 relative to $\boldsymbol{k}^*_{null}$ and commit to more total capacity (dedicated plus flexible) that can be used for node 2.*

$$\psi^I_1 < k^*_{null;1} \ \ \boldsymbol{and} \ \ \psi^I_2 + \psi^I_3 > k^*_{null;2} + k^*_{null;3}.$$

  *(4) In region $\hat{\Omega}_{II}$ (where $d_1$ is high and $d_2$ is low) with partial demand visibility, it is optimal to commit to less dedicated capacity for node 2 relative to $\boldsymbol{k}^*_{null}$ and commit to more total capacity (dedicated plus flexible) that can be used for node 1.*

$$\psi^{II}_2 < k^*_{null;2} \ \ \boldsymbol{and} \ \ \psi^{II}_1 + \psi^{II}_3 > k^*_{null;1} + k^*_{null;3}.$$

Figure 3.8 confirms that results (1) - (4) of Proposition 8 hold in the adaptation region. In Figure 3.8-(a), the difference in total committed capacity is negative (positive) and decreases (increases) as observed demands $d_1$ and $d_2$ decrease (increase) in the adaptation region. In Figure 3.8-(b) and (c), the same pattern is observed in the adaptation region for the difference in total capacity to be used by nodes 1 and 2, respectively. Thus, when observed demands at the two nodes are high (low) in the adaptation region, the decision maker decreases (increases) total capacity (dedicated + flexible) and the capacity available at each node.

In summary, we have shown that optimal capacity commitments in a network with partial demand visibility and flexible capacity are different from the null information

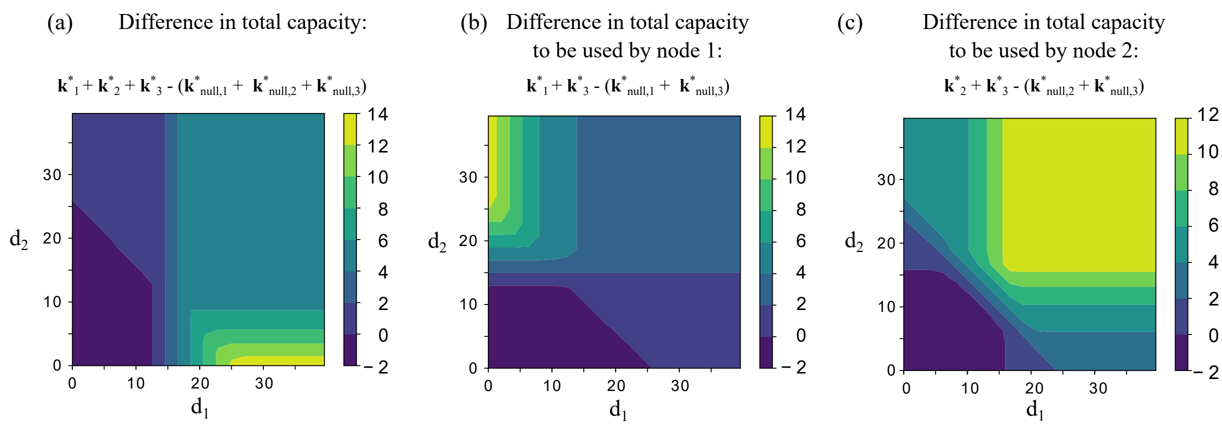


Figure 3.8. (a) Difference in total committed capacity relative to the null information solution. (b) Difference in total committed capacity to be used by node 1 relative to the null information solution. (c) Difference in total committed capacity to be used by node 2 relative to the null information solution.

scenario because partial demand visibility makes it optimal to adapt to observed demand within bounds. Within the adaptation region, when observed demand at one node (say $d_1$) is low and at the other node $d_2$ is high or vice versa, commitment to dedicated capacity decreases and commitment to flexible capacity increases.

### 3.5. Flexible Capacity Can Increase the Value of Demand Visibility

Our goal in this section is to understand how the availability of flexible capacity impacts the value of partial demand visibility in the presence of temporal commitment. Whereas past literature has generally found flexibility and partial demand visibility to be substitutes, we show that the two mechanisms can act both as substitutes and complements in the presence of commitment, leading to the insights first presented in Figure 3.2.

### 3.5.1 Demand Visibility Can Increase the Commitment to Flexible Capacity

In this section we show how the optimal commitment to flexible capacity changes with visible demand. (In Appendix B.2 we discuss how the optimal commitments to dedicated capacity change with visible demand.) We first show that if flexible capacity is not too expensive and the observed demand at one node is high while that at the other is low with a large difference between the two observed demands (we refer to this scenario as demands with large difference), the optimal committed flexible capacity is higher with demand visibility than in the null information case.

**Theorem 4.** ***[The Impact of Demand Visibility on Flexible Commitments]*** *Assume that the cost of committed flexible capacity is sufficiently small ($c_3 - c_1 > 0$ is small enough). If observed demand $d_1$ is low (high) enough and $d_2$ is high (low) enough, the optimal commitment to flexible capacity is larger than the optimal commitment under null information, i.e., $\psi_3^I > k^*_{null;3}$ and $\psi_3^{II} > k^*_{null;3}$.*

Theorem 4 shows that partial demand visibility increases the commitment to flexible capacity when the visible demand at one node is low and the other is high. Without demand visibility, flexible capacity is exclusively used to pool demand (e.g., [**51**] and [**128**]). With partial demand visibility, when the observed demand of one node is high and that of the other node is low, flexible capacity both adapts to visible demand and pools future demand that is not visible.

Figure 3.9 numerically validates the results of Theorem 4. The lines in Figure 3.9-(a) represent levels of constant total demand (i.e., $d_1 + d_2 = \theta \equiv$ constant). As we move away from the center ($d_1 = d_2$) along each line, the difference in observed demand

$|d_1 - d_2|$ increases. We represent this direction of change in Figure 3.9-(b) using the difference $d_1 - d_2$. Figure 3.9-(b) shows that it is optimal to commit to higher levels of flexible capacity as the difference in observed demand increases. For a large difference in observed demand, committing to flexible capacity is more attractive because it can be used to serve the current high demand node while still providing the flexibility to react to future demand fluctuations during the commitment horizon by serving either node.

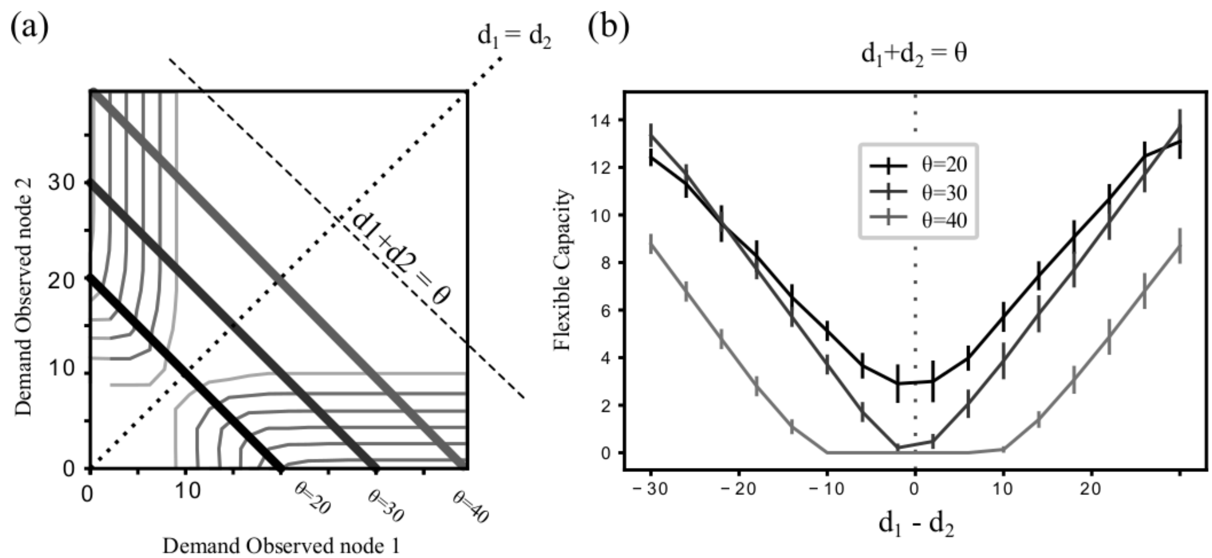


Figure 3.9. Flexible capacity under demand visibility. (a) Lines of fixed total demand ($d_1+d_2 = \theta \equiv$ constant). (b) Commitment to flexible capacity as a function of $d_1 - d_2$.

Adapting the committed flexible capacity for each commitment horizon to observed demand impacts the average optimal commitment to flexible capacity. Figure 3.10-(a) shows the average optimal committed flexible capacity as the cost of flexible capacity changes for both the partial demand information and null information cases. When the cost of flexible capacity is very expensive, i.e., $c_3 \geq c_1 + c_2$ (Region V in Figure 3.10), the average optimal commitment to flexible capacity is 0 in both the partial information and null information cases. As the cost of flexible capacity decreases, the average commitment to flexible capacity increases for both the partial information and null information cases. For costs of flexible capacity between about 0.47 and 0.6 (highlighted in grey in Figure 3.10-(a)), observe that the average commitment to flexible capacity under partial visibility is larger than under null information. In other words, in this region, demand visibility increases the average commitment to flexible capacity. We formalize this observation in Theorem 5 and show that this region always exists.

**Theorem 5.** *[**Average Commitment to Flexible Capacity**] There always exist flexible capacity costs $c_3 \in [c_3^*, c_1+c_2]$ where average commitment to flexible capacity under partial demand visibility is at least as large as the commitment to flexible capacity under null information, $\mathbb{E}_{\mathbf{D}}[k_3^*(\mathbf{D})] \geq k^*_{null;3}$. Under mild conditions, the inequality holds strictly.*

The following section quantifies the benefit of partial demand visibility and flexible capacity using the value of information.

### 3.5.2 The Value of Partial Visibility in the Presence of Flexible Capacity

In this section we quantify the benefits of demand visibility when signing commitment contracts in the presence of flexible capacity by comparing the reduction in costs with respect to the null information scenario to obtain the *Value of Information* (see Definition 5) similar to [**30**]. We use the *Value of Information without Flexible Capacity* as a benchmark to quantify the value of information when the system is analyzed as two independent single-node networks in the absence of flexible capacity (as in Section 3.4.2).

**Definition 5.**

- *We define Absolute Value of Information as the per-period expected cost reduction from partial information:* $AVoI = V(\mathbf{k}_{null}) - \frac{1}{\tau}\mathbb{E}_{\mathbf{D}}[Cost_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D})]$.
- *We define Relative Value of Information as the expected cost reduction with partial information relative to the null information case:* $RVoI = 100 \times \frac{AVoI}{V(\mathbf{k}_{null})})$.
- *We define the (Absolute) Value of Information of a Dedicated Network,* $VoI_{\mathbb{D}}$ *(*$AVoI_{\mathbb{D}}$*), as the* $VoI$ *(*$AVoI$*) when only dedicated capacities are used.*

We use AVoI as the objective when proving our analytical results. However, to make the figures scale free, the results of our computational experiments are shown using RVoI.

Following [**34**], we evaluate the impact of flexibility on the value of partial demand visibility by changing the cost of flexible capacity from high to low. Figure 3.10 shows the average commitment to flexible capacity and RVoI for the null and the partial information solutions as the cost of the flexible capacity changes.

Figure 3.10-(a) shows that it is optimal to commit to no flexible capacity (for both the null and partial information cases) when the cost of flexible capacity is very high (Region V). Figure 3.10-(b) shows the RVoI from partial demand visibility with only a dedicated network (no flexible capacity used). As the cost of flexible capacity decreases below a threshold (0.6 in Figure 3.10), the average commitment to flexible commitment in the partial information case is greater than the commitment to flexible capacity in the null case (see the shaded portion in Region C in Figure 3.10-(a)). As the cost of flexible capacity decreases, the average commitment to flexible capacity increases for both the partial information and null cases. Observe the cost of flexible capacity (roughly 0.47) where the average commitment to flexible capacity in the partial information case equals the average commitment to flexible capacity in the null information case. This is also the

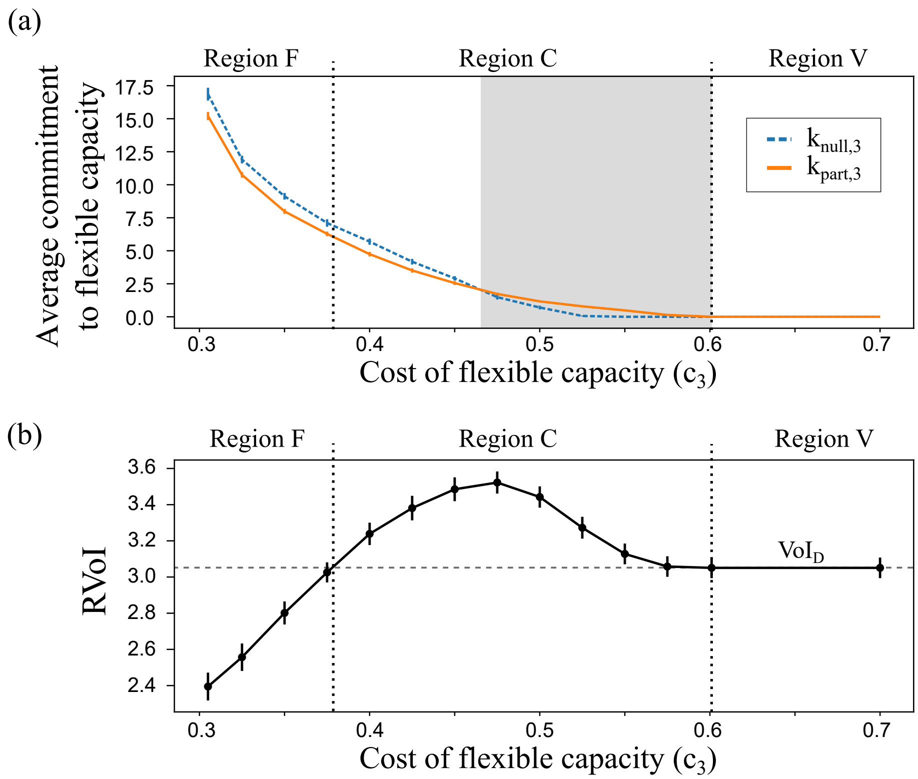


Figure 3.10. VoI as a function of flexible capacity cost $c_3$. (a) The average commitment to flexible capacity with null and partial information. (b) RVoI.

point where the RVoI is maximized. As the cost of flexible capacity decreases below 0.47, the average commitment to flexible capacity in the partial information case drops below the average commitment to flexible capacity in the null case, and RVoI decreases.

We define Region F as the range of costs of flexible capacity where $AVoI < AVoI_D$. In region F, the commitment to flexible capacity is large enough to hedge against almost all future demand uncertainty, thus decreasing the value of partial demand visibility. In Region F, flexibility and partial demand visibility behave like substitutes as observed in [**77**] and [**15**].

Observe that in Region C, between Regions F and V, it is optimal to commit to both flexible and dedicated capacity, and the value of information is the highest. We formally define Region C below.

**Definition 6.** *Region C is the set of flexible capacity costs where* $AVoI \geq AVoI_{\mathbb{D}}$*. Region C corresponds to the flexibility cost interval* $[\hat{c}_3, c_1 + c_2)$*, where* $\hat{c}_3 = \min_{AVoI \geq AVoI_{\mathbb{D}}} c_3$*.*

Flexible capacity and demand visibility complement each other in Region C where the optimal commitment to flexible capacity changes with observed demand, thus increasing the value of partial demand visibility compared to Region V, where no flexible capacity is used. In Figure 3.10 observe that the value of information is the highest when the average commitment to flexible capacity with partial demand visibility equals the commitment to flexible capacity in the null information case. To show that this is true in general, we first show that the gradient of the value of information curve relative to the cost of flexible capacity equals the difference in average commitment to flexible capacity with null information and with partial demand visibility (the distance between the solid orange and dashed blue lines in 3.10-(a)).

**Theorem 6.** *[**Marginal Difference**] The gradient of AVoI with respect to* $c_3$*, the cost of flexible capacity, is equal to the difference of the expected commitment to flexible capacity under null information and under partial visibility, i.e.,* $\frac{\partial AVoI}{\partial c_3} = k^*_{null;3} - \mathbb{E}_{\mathbf{D}}[k^*_3(\mathbf{D})]$*.*

From Theorem 5, at the transition from Region V to Region C, the average commitment to flexible capacity with partial information is greater than the zero commitment under null information. From Theorem 6 it thus follows that the gradient of AVoI is negative at the transition from region V to region C. In other words, at the transition from region V to region C, the value of information increases as the cost of flexible capacity decreases.

In contrast, we transition from Region C to Region F when the cost of flexible capacity is very low. As a result, it is optimal to commit to a large amount of flexible capacity in the null scenario. At the transition from region C to region F the expected commitment to flexible capacity under partial information is smaller than the null case. This is because the visible demand is unlikely to have a large enough difference in demand between the two nodes to require a higher level of flexible capacity than the null case. In contrast, it is more likely to observe demands at the two nodes that are not different enough to justify committing as much flexible capacity as the null case. In other words, there is a higher probability that demand visibility will result in a lower level of commitment to flexible capacity than the null scenario. By Theorem 5, AVoI thus increases as the cost of flexible capacity increases at the transition from region F to region C. Given that the gradient of AVoI is positive at the transition from region F to region C and negative at the transition from region C to region V, there must be an intermediate point where the gradient is zero. This point maximizes the AVoI and occurs where the average commitment to flexibility under partial information equals the commitment to flexibility in the null scenario, as shown in Figure 3.10. The AVoI is higher with partial information because the decision maker either commits to more or less flexibility than the null scenario based on visible demand.

We show in Theorem 7 that Region C always exists and that under mild conditions, the VoI in Region C is strictly greater than the $VoI_{\mathbb{D}}$ in Region V. Therefore, the maximum AVoI is achieved in Region C.

**Theorem 7.** ***[The Value of Information and Flexible Capacity]*** *There always exists a band of flexible capacity costs* $[\tilde{c}_3, c_1 + c_2]$ *(Region C), where the value of information with flexibility is at least as large as the value of information without flexibility,* $AVoI \geq AVoI_{\mathbb{D}}$*. Under mild conditions, the inequality holds strictly. Hence, there always exists a band where flexible capacity and partial demand visibility complement each other.*

As shown in Theorems 2, 4, 5, demand visibility creates value in Region C by allowing the decision maker to adapt the commitment to flexible capacity based on the observed demand. In Region C, using demand visibility to adapt to observed demand increases the marginal value of flexible capacity, and vice versa. Therefore, partial demand visibility and flexible capacity behave like complements in Region C.

To conclude, we find that there is a range of flexible capacity cost where the value of partial demand visibility is higher than in the absence of flexible capacity, i.e., flexibility and partial demand visibility behave as complements. In the next section, we evaluate the impact of commitment duration, correlation and uncertainty on the complementary behavior of flexibility and partial demand visibility.

## 3.6. Sensitivity Analysis

In this section we explore analytically and numerically how the complementarity of demand visibility and flexible capacity changes with key parameters of the model such as demand uncertainty, demand correlation and commitment duration. To obtain analytical insights, we set up a simplified model that preserves the key features of our general model. We then conduct numerical experiments to confirm that the analytical insights from the simplified model hold in the general case as well.

### 3.6.1 A Simplified Low-High Model

To obtain sensitivity analysis insights analytically, we consider a simplified version of our model where demand at each node has only two possible realizations: Low ($d_l = \bar{d} - \delta$) or high ($d_h = \bar{d} + \delta$). Hence, demand at both nodes $\boldsymbol{D}$ can have one of four possible realizations: $\boldsymbol{d} \in \{(d_l, d_l)', (d_l, d_h)', (d_h, d_l)', (d_h, d_h)'\}$, which occur with probability $p_{ll}, p_{lh}, p_{hl}$ and $p_{hh}$, respectively. For simplicity, we assume that the two nodes are symmetric (i.e., $p_{lh} = p_{hl}$ and $c_1 = c_2$). It is easy to see that there are only five possible commitment vectors for such a model: $k_a = (d_l, d_l, 0)'$, $k_b = (d_h, d_l, 0)'$, $k_c = (d_l, d_h, 0)'$, $k_d = (d_h, d_h, 0)'$, $k_e = (d_l, d_l, d_h - d_l)'$. We denote this model as the *L-H Model*.

Theorem 8 uses the L-H model to obtain analytical results that we subsequently use to study the impact of commitment duration, uncertainty and correlation on the value of information in the presence of flexible capacity. We assume that the commitment duration is sufficiently large, i.e. $\tau$ is not close to 1.

**Theorem 8** (**L-H Model**). *Under the L-H Model, where the commitment duration is sufficiently large (not too close to 1):*

- ***(1)*** *There exists a region where partial visibility and flexible capacity behave like complements (A strict version of Theorem 7).*
- ***(2)*** *The maximum absolute value of information is obtained as $AVoI^*(\delta, \tau) = \delta A_p(\tau)$, where $A_p(\tau)$ is a function independent of $\delta$ and decreasing in $\tau$.*
- ***(3)*** *The width of Region C, where the two mechanisms behave like complements, is an increasing function in $\tau$ and independent of $\delta$.*

*(4) The maximum value of information is obtained at* $c_3 = \min\{c_1 + c_2 - c_s p_{hh}, c_s(1 - p_{ll})\}$*, where the marginal value of using flexible capacity under null information is 0.*

We use the results of Theorem 8 to understand the impact of commitment duration, demand uncertainty and demand correlation on the value of information.

### 3.6.2 The Impact of Commitment Duration on Value of Information

Consistent with the general model, Theorem 8-(2) shows that the maximum value of information decreases as the commitment duration $\tau$ increases. Theorem 8-(3) shows that as $\tau$ increases, demand visibility and flexible capacity behave like complements for a wider range of flexibility cost, i.e., region C expands.

Figure 3.11 validates these results numerically for the general model where demand follows a continuous gamma distribution. Observe that the maximum value of information decreases as the commitment duration $\tau$ increases. Consistent with Theorem 8-(3), the width of region C, where flexible capacity and partial information are complements, expands as $\tau$ increases.

### 3.6.3 The Impact of Demand Variability on Value of Information

Consistent with previous literature (e.g. [**19**]), in this section we show that the maximum value of information increases with demand variability.

For the LH-Model, increasing the value of $\delta$ increases the variance of demand at both nodes. Theorem 8-(2) shows that the AVoI increases linearly with the demand variability

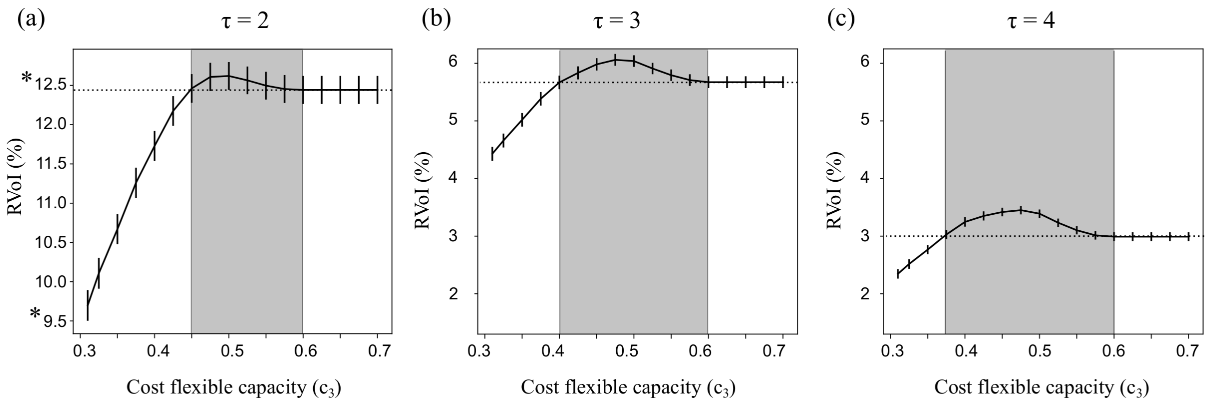


Figure 3.11. VoI for different commitment duration, $\tau$, with bivariate Gamma distribution and $\mathbf{c} = (0.3, 0.3, c_3)$. Asterisks highlight that the y-axis scale is different in (a).

parameter, $\delta$. Theorem 8-(1), (2), (4) show that the existence of Region C and the position of the maximum AVoI is independent of the demand variability parameter $\delta$.

Figure 3.12 shows the RVoI for the general model using a gamma distribution when the coefficient of variation of demand increases from 0.3 to 0.7. Observe that the value of information increases with the coefficient of variation. The shape and width of Region C, however, remains unchanged, consistent with the prediction of the L-H model.

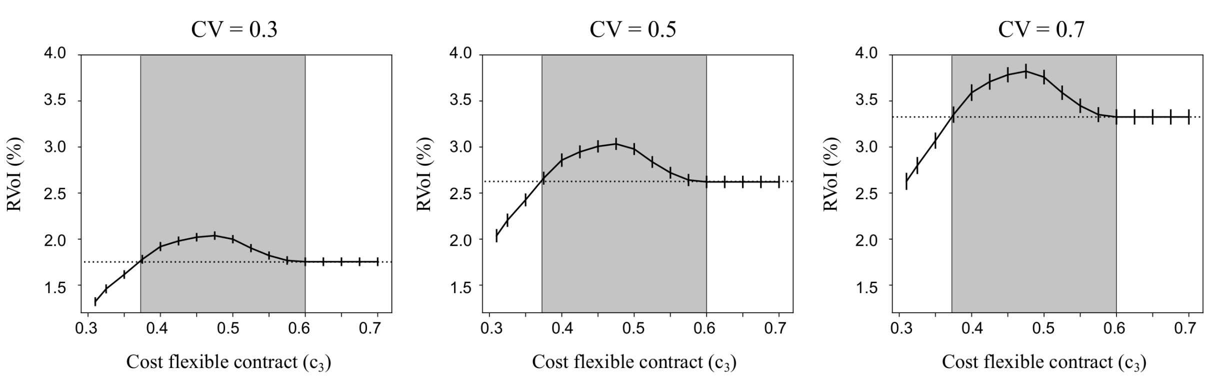


Figure 3.12. VoI for different coefficient of variation, with bivariate Gamma distribution and $\mathbf{c} = (0.3, 0.3, c_3)$.

### 3.6.4 The Impact of Correlation on the Value of Information

In this section we show that the cost of flexibility at which the value of information is maximized decreases as the correlation of demand between the two nodes increases.

We have shown in Theorem 6 that the value of information achieves its maximum when the marginal value of flexibility is 0. For the LH-model, that point is given by $c_3^* = \min\{c_1 + c_2 - c_s p_{hh}, c_s(1 - p_{ll})\}$. As correlation of demand at the two nodes increases, $p_{ll}$ and $p_{hh}$ increase, and $c_3^*$ decreases.

We validate this finding numerically for the general model in Figure 3.13. Observe that when correlation of demand at the two nodes increases, the maximum value of information is achieved for lower cost of flexibility. Also observe that increasing demand correlation expands the range of cost of flexible capacity where flexibility and demand visibility are complements.

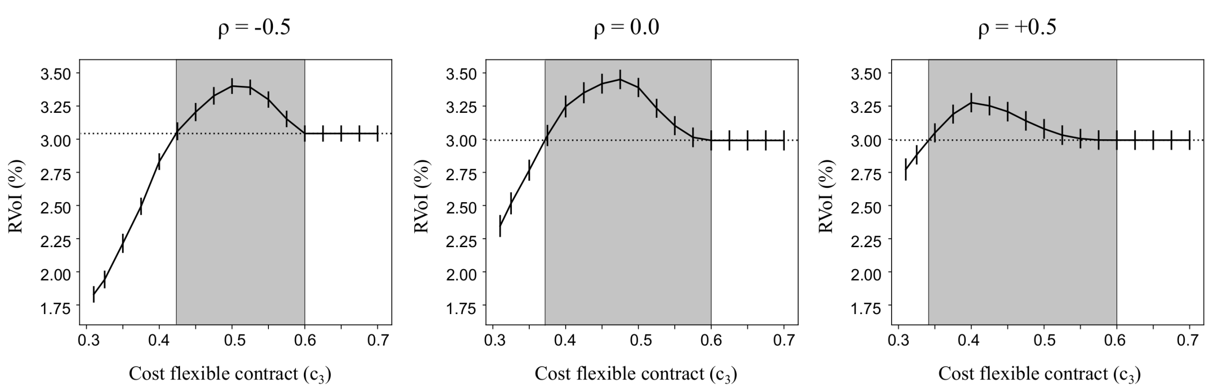


Figure 3.13. VoI for different correlation coefficient with bivariate Gamma distribution and $\mathbf{c} = (0.3, 0.3, c_3)$. Region C (shaded region) increases with correlation.

## 3.7. The Impact of Seasonal Demand Surges

In this section, we relax the IID assumption on demand and study the impact of seasonality and prediction uncertainty. In Section 3.7, we show that combining flexible capacity and demand visibility is particularly valuable under demand surges. In Section 3.7, we show that, under seasonal surges with prediction uncertainty, the two hedging mechanisms further complement each other, amplifying the results in the previous sections.

### The Impact of Demand Seasonality

Demand for the furniture manufacturer (see [**19**]) shows seasonal demand fluctuations (e.g. a surge in demand for the President's Day sale). We extend the previous results to the case where demand in the commitment horizon follows seasonal trends.

In Figure 3.14, we numerically explore the impact of seasonality on the interaction between flexible capacity and demand visibility with commitment. We fix the commitment horizon $\tau = 4$ and compare how the VoI changes with the cost of flexible capacity under three representative situations: Observed demand in the visible period is an expected seasonal *surge* with $\mathbb{E}[D_i] = 20$ that decreases linearly to $\mathbb{E}[D_i] = 10$ in the commitment horizon; *IID* where expected demand is stable at $\mathbb{E}[D_i] = 15$ across the commitment horizon; and *drop* where the observed demand in the visible period is expected to be low with $\mathbb{E}[D_i] = 10$ increasing linearly to $\mathbb{E}[D_i] = 20$ over the commitment horizon.

Observe that the $VoI$ is highest and Region C (where flexible capacity and demand visibility behave like complements) is the widest when demand in the information horizon corresponds to a seasonal surge. Demand visibility is most valuable when the visible

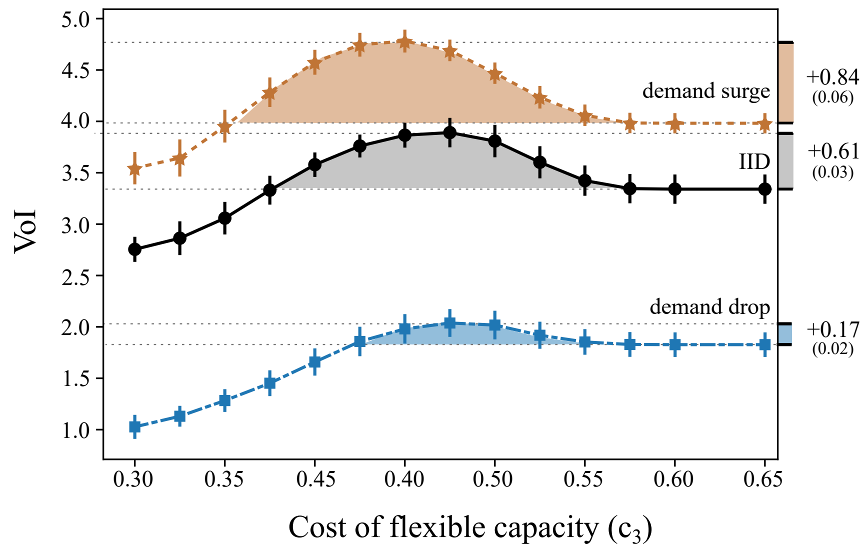


Figure 3.14. VoI as a function of flexible capacity cost $c_3$ under different seasonal trends. From top to bottom: demand surge, stationary (IID) and demand drop. Numbers in parenthesis report the standard error.

period includes a surge. As a result, the findings in Section 5 are amplified when a demand surge is visible over the commitment horizon.

**Prediction Uncertainty in Seasonal Surges**

Demand at the furniture company displays seasonality given the periods when the product is discounted (such as Labor Day). While the furniture company expects demand to surge during the sale period, predicting the extent of the surge is quite challenging. As a result, they recalibrate their prediction model as demand is observed. There is plenty of literature (e.g. [**103**] and [**5**]) indicating that the value of demand visibility increases with prediction uncertainty. In this section we explore the interaction between partial demand visibility and flexible capacity under prediction uncertainty during seasonal surges.

We use a stylized Bayesian model in which demand at both nodes could either come from a distribution $f_h$ with high expected demand $\boldsymbol{\mu}_h$ (high surge) or a distribution $f_l$ with low expected demand $\boldsymbol{\mu}_l$ (low surge) with $\boldsymbol{\mu}_h \succ \boldsymbol{\mu}_l$. Starting with a prior, the decision maker updates her beliefs based on observed demand as follows:

- **•:** Prior to observing the next period demand in period $t$, the decision maker has a prior probability $p$ that demand $\mathbf{D}$ follows distribution $f_h$ with high expected demand. The prior probability $p$ captures the amount of prediction uncertainty. If $p = 1$ ($p = 0$), the decision maker is certain that demand follows a high (low) demand surge with distribution $f_h$ ($f_l$). If $p = 1/2$, demand is equally likely to follow high or low demand distributions and, therefore, prediction uncertainty is the highest.
- **•:** After observing the next period demand $\mathbf{d}$ the decision maker updates the posterior probability $p'(\mathbf{d})$ according to Bayes' rule: $p'(\mathbf{d}) = \mathbb{P}(\mathbf{D} \sim f_h|\mathbf{d}) = \frac{pf_h(\mathbf{d})}{pf_h(\mathbf{d})+(1-p)f_l(\mathbf{d})}$.
- **•:** After obtaining the posterior probability, the decision maker commits to capacities $k_1$, $k_2$ and $k_3$ for a commitment horizon of $\tau$ periods.

In Figures 3.15 - 3.17 we numerically study the impact of prediction uncertainty on capacity commitments and the value of demand visibility. While results hold in more general cases, we assume that, conditioning on the distribution being known, demand at the two nodes is identically distributed and independent. In our numerical experiments, $f_h$ and $f_l$ are multivariate Gamma distributions with $\boldsymbol{\mu}_h = (14, 14)'$ and $\boldsymbol{\mu}_l = (8, 8)'$ and the same variance $\sigma = 5^2$. The prior probability that the distribution is $f_h$ is $p = 0.75$. Figure 3.15-(a) shows samples of observed demand $\boldsymbol{d}$ of $f_h$ and $f_l$. Figure 3.15-(b) shows the posterior update $p'$ after the next period demand is observed.

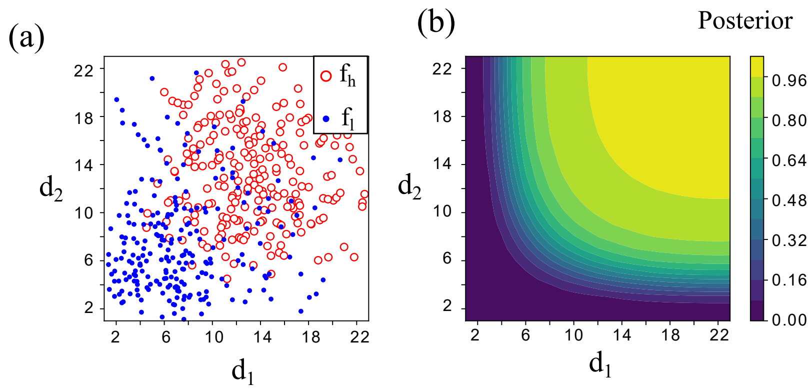


Figure 3.15. (a) Samples of $f_h$ and $f_l$ and (b) posterior.

Our goal in this section is primarily to study the impact of prediction uncertainty. While duration of demand surges is typically short (few weeks), for the purpose of illustration, in Figure 3.16 we begin with isolating the effect of prediction uncertainty by using a long commitment horizon ($\tau \to \infty$) and equally long demand surge duration, a situation for which adaption is less valuable.

Figures 3.16-(a) and (b) show how commitment to dedicated capacity changes with observed demand. The red and blue arrows in the legend show the capacity commitment when the distribution is known with probability 1 to be $f_h$ or $f_l$ respectively. When observed demand at both nodes is high, e.g. $d_1 = d_2 = 14$, the decision maker obtains a high posterior probability close to 1 (see Figure 3.15-(b)) and is certain that demand during the surge is high. As a result, commitment to dedicated capacity is high (around 12.5). The converse occurs (posterior probability close to 0) when observed next period demand is low, e.g. $d_1 = d_2 = 4$. Observe that there exist a region around $d_1 = d_2 = 7$, with observed demand not too high or low, where the posterior is around $p' = 1/2$, i.e., prediction uncertainty is the highest. In that case, the decision maker commits to moderate levels of dedicated capacity (roughly $k_1 = k_2 = 9.5$) that is between the commitment made when demand follows $f_h$ (red arrows at $k_1 = k_2 = 12.5$) or $f_l$ (blue arrows at $k_1 = k_2 = 6.5$).

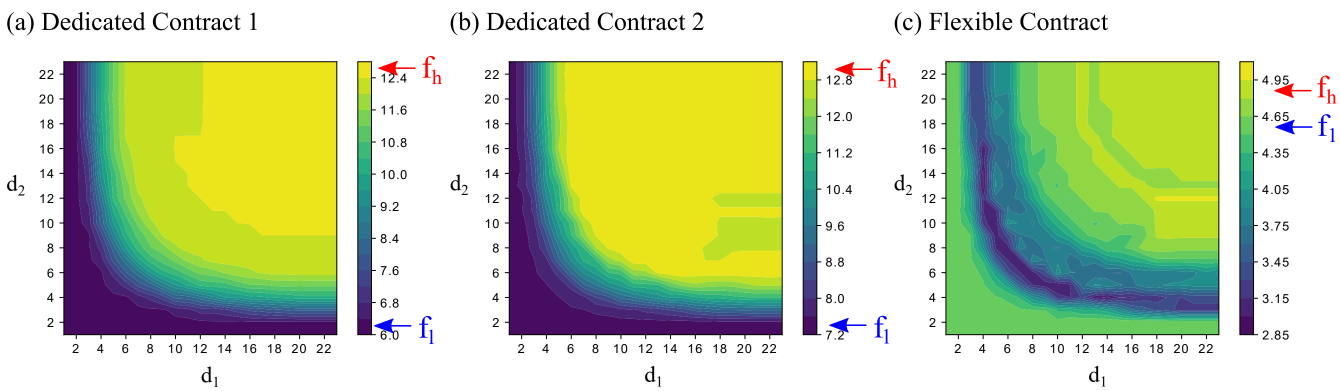


Figure 3.16. Capacity commitment with prediction uncertainty for the seasonal surge (commitment horizon $\tau \to \infty$). Capacity costs are $\mathbf{c} = (0.3, 0.3, 0.4)'$.

Figure 3.16-(c) shows the corresponding commitment to flexible capacity. When observed demand at both nodes is high ($d_1 = d_2 = 14$) or low ($d_1 = d_2 = 4$), the posterior is close to $p' = 1$ or $p' = 0$. In either case the decision maker commits to a flexible capacity level similar to the commitment under $f_h$ or $f_l$ (roughly $k_3$ is between $4.6 - 4.9$), because $f_h$ and $f_l$ have the same variance.

Now consider the case when observed demand is around $d_1 = d_2 = 7$, which gives a posterior around 0.5 and has the highest prediction uncertainty. In this case, commitment to flexible capacity is significantly lower (around $k_3 = 2.8$) than the flexible capacity committed under $f_h$ or $f_l$. For this value of observed demand, dedicated capacity commitment is $k_1 = k_2 = 9.5$. Given the high prediction uncertainty, the decision maker is unsure whether demand follows a low or high surge. If demand in future periods follows a low demand surge, $f_l$, the committed dedicated capacity fully satisfies demand at both nodes making further commitment to flexible capacity less valuable. In other words, with high prediction uncertainty, flexible capacity is valuable only if future demand follows a high demand surge, an event occurring with probability 1/2. Consequently, when the posterior prediction uncertainty is high, the decision maker decreases commitment to flexible capacity. Committing to flexible capacity is less valuable when the observed demand increases prediction uncertainty and the posterior is close to 1/2.

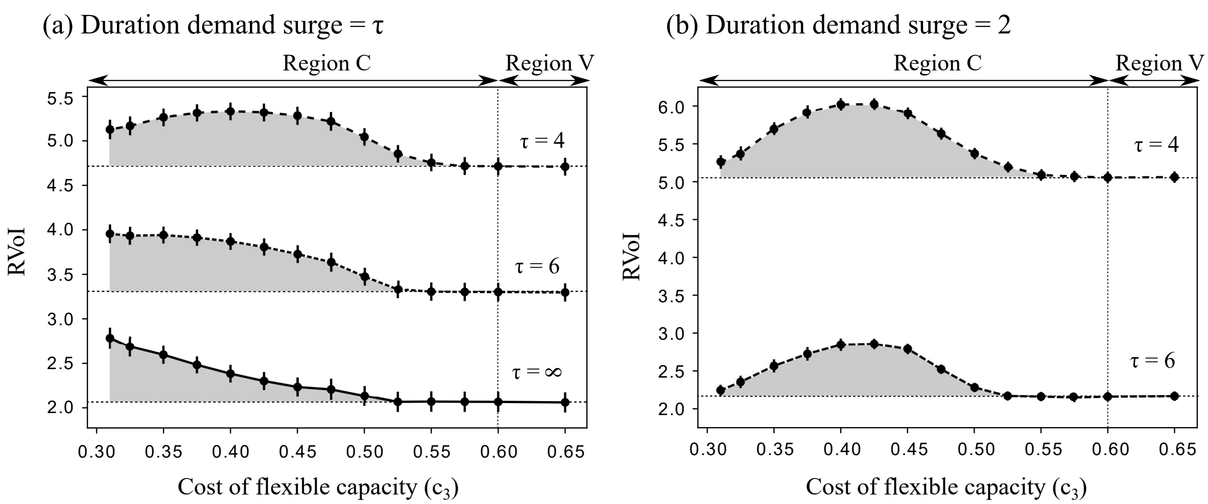


Figure 3.17. VoI with prediction uncertainty for the seasonal surge. (a) Demand surge duration is equal to commitment horizon. (b) Demand surge duration is two periods. After two periods, demand follows $f_l$ with probability 1.

In Figure 3.17, we plot the value of information as the cost of flexible capacity changes. Figure 3.17-(a) shows results when the demand surge lasts for the entire commitment horizon. To isolate the impact of prediction uncertainty, we focus on the solid line corresponding to a long commitment horizon ($\tau \to \infty$). Unlike Figure 3.10, where there exists Region F (where partial demand visibility and flexible capacity are substitutes), there is no such region under prediction uncertainty for surges in Figure 3.17. Observe that the equivalent of Region C of Figure 3.10 in Section 3.5, where flexible capacity and demand visibility behave like complements, extends to the whole range $c_1 < c_3 < c_1 + c_2$. Similar results are obtained when commitment horizon is shorter (dashed lines). VOI is higher in the presence of prediction uncertainty because the decision maker uses demand visibility to counteract prediction uncertainty and adapt capacity commitments to observed demand as described in Section 3.5. Figure 3.17-(b) shows similar results when duration of the demand surge is shorter than the commitment horizon.

In demand surges with prediction uncertainty, the availability of flexible capacity increases the decision makers willingness to acquire demand visibility. In other words, flexible capacity and partial demand visibility behave even more like complements in the presence of prediction uncertainty for seasonal surges. Thus, companies facing prediction uncertainty for seasonal surges that have access to flexible capacity can benefit considerably by being able to acquire partial demand visibility.

## 3.8. Final Remarks

Our work shows that the presence of short term commitment introduces more nuance to the relationship between partial demand visibility and flexible capacity when dealing with demand uncertainty. Whereas the two behave like substitutes when the cost of flexible capacity is very low, they are complements when the cost of flexible capacity is somewhat higher without being too large. As technology creates flexibility and allows short term commitments in a variety of settings such as warehousing, manufacturing and processing, companies will increasingly find it valuable to obtain partial demand visibility.

CHAPTER 4

# Estimation of Customer Availability for Attended Home Delivery

*Chapters 2 and 3 assume that delivery dates to customers' homes are fixed. This chapter explores flexibility in delivery dates and focuses on scheduling policies for last mile distribution, i.e. from spokes to the customer's home. Understanding customers' sensibility to delivery lead time can allow retailers to balance customer satisfaction and shipping costs in attended home-delivery. Using transactional data from the furniture company, we find that, when customers choose their preferred delivery date, speed of delivery is of limited importance relative to other priorities and customers can be easily segmented based on day of week preferences.*

## 4.1. Introduction

Last mile distribution is one of the most challenging and expensive operations of the retail supply chain [**10**]. It is estimated that 53% of the total cost of transportation of home delivery is attributed to the last mile delivery distribution [**44**]. The importance of the last mile delivery challenge has rapidly increased in the last decade with the emergence of e-commerce [**94**]. More recently, the COVID-19 outbreak has accelerated the demand for home delivery services [**10**]. In addition to the operational challenges, increasing customers' expectations and the competing pressure of companies such as Amazon, Alibaba or Walmart have pushed retailers to assume the entire cost of transportation [**98, 70**], including the expensive last mile delivery portion, as well as to enhance the customer experience by increasing speed of delivery. Such an effort usually comes at high costs for companies and the environment [**42, 104, 37**]. Hence, reducing the cost of last mile distribution while meeting customer expectations has become crucial.

In this chapter we focus on attended home delivery, where customers must be present at the time of delivery, a type of delivery that is commonly adopted in several sectors such as e-groceries [**2**], delivery of health care services [**84**], telecommunications services (such as AT&T and Xfinity) [**23**], delivery of medical drugs and office supplies [**28**], or furniture [**19**]. Companies engaged with attended home delivery face a trade off between customer satisfaction and shipping costs. Offering a cost-efficient but limited menu of available dates can negatively impact customer experience. On the other hand, letting customers choose any date regardless the transportation capacity can be quite costly for the companies' home delivery teams. It is in the companies' interest to quantify and understand customer preferences for home delivery to balance customer satisfaction and delivery costs.

To that end, we explore the transactional data from 271 retail outlets owned by the furniture company that motivated our work in Chapters 2 and 3. When sales are finalized at the retail outlets, customers choose a delivery date among an offered menu of alternatives. Dates are offered based on the remaining capacity of delivery routes of each region. The sequence of each route and the final delivery time are determined later, two days before the actual delivery date, and customers are then notified of the delivery time window (e.g. from 9 a.m. to 11 a.m.). Our data set contains information about the sales transactions and the last mile delivery operations during the season 2016-2017, which we use to derive customer priorities when choosing attended home delivery dates, and identify managerial insights to balance customer satisfaction and shipping costs.

We use a behavioral model that captures the interaction between customers and the retail company when customers choose attended home delivery dates. When choosing attended home delivery, customer's availability to be at home at the time of delivery plays a fundamental role. To the best of our knowledge, ours is one of the first studies that propose a sequential logistic approach to model customer choice. Our model allows us to identify the key factors driving customer choice, and to estimate how likely it is that a customer chooses a specific day available for attended home delivery (i.e. probability that staying at home on a specific day is convenient for a customer).

The data validate insights from the company that the days immediately after the purchase date are usually not available for the customer. For the home delivery team, there may be an opportunity to slightly extend fulfillment lead times beyond the first four to five days, and more efficiently utilize last mile capacity. One week after the in-store purchase date, customer availability significantly increases at a rate of $1.5(\pm 0.1)$ % per each additional day since the in-store purchase date.

We find that the day of week has a strong impact on customers' availability. While this result is not surprising, our approach allows us to explore, quantify and model how the effect of day of week varies across different customer segments. Our model, which is built around this intuition, provides simple ways to segment customers when they visit the store to make a purchase. We observe that customers that visited the store and made a purchase on Tuesdays, Wednesdays, and Thursdays have high availability during weekdays. Customers that visited the store and made a purchase on Saturdays and Sundays, have, on average, low availability during weekdays from Tuesdays to Thursdays, and more frequently schedule the attended home delivery the first available Friday or Saturday. For the furniture company, saving some delivery capacity, particularly on Fridays and Saturdays, can significantly decrease waiting times of customers that make a purchase on Saturdays and Sundays, while minimally affecting customer segments with higher availability. In general, we find that the day of week that customers visited the store and made a purchase is a strong predictor of the customers' availability. Our results show that the day of week of purchase is an intuitive and simple way to segment customers. This result suggests that, with easy-to-implement customization of delivery options, the company can potentially meet customer's needs while more efficiently deploy last mile transportation capacity.

Several studies have shown that incorporating customer behavior in scheduling operations can be highly valuable (e.g., [**65**], [**48**]). For the furniture company, our model can be used to estimate how potential adjustments of delivery capacity might ultimately impact customer experience.

This chapter is organized as follows. Section 4.2 describes the managerial context that motivates our work, states the research questions, and reviews the related literature. Section 4.3 describes the data and the behavioral model that we use to understand customer choice of attended home delivery. In Section 4.4, we discuss the methods and the tools that we use in our study. Section 4.5 presents our key empirical findings. We discuss our results and describe managerial insights in Section 4.6. We provide final remarks in Section 4.7.

## 4.2. Context, Related Literature and Hypotheses

In Sections 4.2.1 and 4.2.3 we describe the decision process faced by a retail company and state the research questions that motivate our work. In Section 4.2.2, we discuss the literature that is related to our work.

### 4.2.1 Context

Our work is motivated by the home delivery scheduling operations of a major furniture retail company. The company's most popular products are high-quality pieces of furniture with prices ranging from $1,000-$5,000, sold in more than 500 outlets owned by the retail company. The principal role of the retail outlets is to display the popular products in showrooms where customers can have a first-hand experience before committing to purchase. The retail company operates with a make-to-order model: once a sale is closed at the retail outlets, the retail company commits to deliver the purchased products at the customer's home on a date chosen by the customer. Once an order is entered into the system, products are prepared and distributed from a central warehouse to the distribution centers, from where the last mile delivery routes depart.

The last mile distribution and product installation are some of the most challenging and costly operations for the supply chain team of the retail company. Due to the large size of the deliveries, and since most products need to be installed by technicians, customers are asked to be present at the time of delivery, an operation known as Attended Home Delivery (AHD). The last mile delivery is operated by home delivery teams of drivers and technicians that are employed by the furniture company. Two days before a scheduled delivery, the home delivery team calls the customer receiving the delivery to agree upon a time window, i.e., the estimated time arrival at the customer's home. Therefore, when customers choose a candidate day for delivery, they need to evaluate whether such a date is available or convenient in their personal schedule.

Customer satisfaction is highly considered by the retail company, which targets customer segments that value its luxury products and, therefore, expect a customer experience of equal quality. To that aim, the retail company assumes the cost of transportation (free shipping) and, to ensure that delivery lead times are met, delivery capacity is allocated in advance (i.e., number of delivery teams). When planning delivery capacity, the retail company faces a trade-off between customer satisfaction and the cost of transportation. Ample delivery capacity shortens delivery lead times and allows the retail company to offer a wide range of candidate dates for delivery. However, excess delivery capacity results in idle delivery teams and unutilized costly resources. On the other hand, tight capacity increases resource utilization at the cost of offering more limited menus of candidate days that usually have longer delivery lead times, which might, potentially, harm customer experience.

Our work focuses on the customer decision process of choosing a date for AHD. When a sale is closed, a sales representative and the customer interact to agree upon an available date among the dates with remaining delivery capacity. It is customary for sales representatives to suggest the earliest day available for delivery, and ask for the customer's availability. If the first available date suggested is not convenient for the customer, future dates are offered until a date is found to be convenient. Therefore, dates are usually sequentially offered from earlier to later dates.

Note that because of the way dates are typically offered to the customer, the scheduled date tends to be the first that is available for the customer among the dates that are available for the retail company. The final dates chosen by the customers result from the interaction of two parties with different availability: First, the retail company, whose delivery teams have limited capacity and, therefore, only dates that are available for delivery are offered; and, second, the customer, who must be present at the time of delivery and will only choose a delivery date provided that such date is convenient in his/her personal schedule.

A big body of the literature has focused on shortening the delivery lead time to offer very fast deliveries [**1, 70, 104**]. However, in the context of AHD, only those dates that are available in the customers' schedule may be chosen for delivery. In this context, speed of delivery is relegated to a secondary role below customer convenience. In this case, it is not clear that the same intuition on the value of speed of delivery (short delivery lead times) applies.

With the objective of gaining insight into the relative importance of speed of delivery, the supply chain team of the retail company asked the following questions: "What are the benefits of decreasing or increasing delivery lead time? What is more relevant for the customers, day popularity or speed of delivery (i.e. availability of short lead times)?" "Can the company reduce costs by decreasing delivery capacity without impacting customer satisfaction?"

To answer these questions, we use a transactional data set from the company, and we model the customers' decision process when scheduling AHD. Our work is related to the retail and delivery literature, as well as to the behavioral scheduling literature. In the following section we review the related literature.

### 4.2.2 Related Literature

With increasing competition and customers' expectations in the retail sector, retailers are looking for ways to improve delivery, introducing functionalities such as ship-to-store (STS) [**55**] and redesigning their distribution networks to get closer to the customers [**57**]. Accordingly, a large body of the literature has focused on improving customer experience by increasing speed of delivery and offering services such as same-day or two-day home delivery (e.g., [**104**], [**1**], and [**70**]). Several empirical studies have pointed out the benefits of decreasing delivery lead times. [**52**] studies the impact of reducing delivery lead times for a large portion of customers by opening new (closer to customers) distribution centers. The authors find that reducing delivery lead time led to a significant increase of revenues and profits. Studying a natural experiment, [**38**] shows that removing the option of express delivery significantly reduced sales by 14.56% of the Chinese online retailer Alibaba. [**39**] shows that promising one day faster delivery increased profits by 2%. In the omni-channel retail setting, [**82**] shows that the positive impact of increasing delivery lead time is stronger on consumers using online and catalog channels than on consumers using showroom channels.

We extend this work by considering factors that may limit the benefits of delivery speed. Specifically, we focus on AHD, where the customer must be present at the time of delivery. This element of co-creation makes delivery a service that can only be provided when both the company and the customer are available.

Building on research on behavioral scheduling (e.g., [**71**] and [**72**]), we posit that customers may have preferences or availability constraints that affect their choice regarding attended delivery speed. Accordingly, faster delivery may be infeasible or undesirable, which has implications for retailers scheduling transportation capacity. To the best of our knowledge, we are among the first to investigate how and why slower delivery may be preferable under certain conditions in the retail literature.

Outside retail and delivery, there is some research suggesting that increasing speed of service provides little benefits, or can even negatively impact customer experience. In health care scheduling, some papers have focused on the relative importance of speed of service (or speed of access) over other factors that may be ranked as more important for the patients. [**109**] finds that for General Practitioner (GP) scheduling, increasing the proportion of capacity dedicated to same-day appointments decreases patient satisfaction. For most patients, the benefits of speed of access provided by same-day appointments do not compensate the added difficulties in booking appointments in advance. [**108**] designs a discrete choice experiment to study patients accessing primary care. The authors conclude: "Speed of access is of limited importance to patients accessing their GP, and for many is outweighed by choice of GP or convenience of appointment." [**83**] studies patient choice behavior in scheduling medical appointments. The authors design a discrete choice experiment and analyze the answers with a Mixed Logit. They quantify the impact of several factors such as appointment delays (days until the appointment date), in-clinic waiting times, changing doctor assignments, and risk attitudes. Results suggest that health care provides would benefit more from improving all quality measures than just focusing on reducing the appointment delay to zero. These results are along the lines of other findings raising concerns on the relative value of speed of access, e.g.; [**13, 105, 43**]. Our work contributes to the literature by exploring the customers' tradeoffs between speed of service (or delivery), availability, and day preferences. We focus on investigating the reasons by which individual customers might prefer later days over earlier days, and the conflicting priorities of both the customers and the service providers when scheduling services. Our behavioral model reveals that speed of delivery is of limited importance relative to other priorities such as availability and day of the week preferences. As a result, AHD is rarely scheduled on dates immediately after the purchase, even when these dates are offered by the retailer. Customers have strong preference of day of the week, as one can expect. Our analysis quantifies and models these preferences and that the day of the week on which customers physically visit the stores to make a purchase reveals customers' day of the week preferences for AHD. This suggests an easy-to-implement customer segmentation by day of week of purchase that can be leveraged to customize delivery options to customer needs and use more efficiently delivery capacity.

Scheduling AHD and health care appointments share an important similarity: Customers (or users) must schedule and commit in advance, and they must be present at the time the service is delivered. The AHD literature has focused on algorithms and routing policies that incorporate customer behavior in order to offer fast delivery services in the most profitable way [**2, 7, 81, 86, 143, 142**]. Similarly, the incorporation of patient preferences in health care scheduling policies has gained attention in the recent years. [**99**] estimates the impact of waiting time in no-show behavior and find how strategically using waiting time can increase capacity utilization and reduce no shows. [**65**] and [**136**] study scheduling policies for intra-day appointments in which customers have time and practitioner preferences. For inter-day appointments, the authors in [**48**] model customer preferences with a multinomial logit choice model, and study the performance of online and offline scheduling policies. To mitigate the impact of no-shows due to appointments scheduled too far in the future, their results suggest offering all consecutive dates earlier than a threshold date. Our empirical results contributes to this literature by showing that customer availability plays a fundamental role in AHD. We show that customer availability impacts customer choice, it strongly depends on the day of week, and it can be predicted. We find that the company can potentially reduce the cost of the last mile distribution by tailoring delivery capacity and scheduling policies to customer preferences and availability.

In the context of AHD, the following section states a set of hypotheses that we aim to test using the company's data.

### 4.2.3 Theoretical Framework: Hypotheses

We focus on the importance of speed of delivery, or Order-to-Delivery Lead Time (OD-Lead-Time), which is the time in days passed between the day a customer made a purchase and an offered candidate day for AHD. We also study the effect of the day of week on customer choice. In Section 4.6 we show how information about customer priorities can help retailers to deploy costly delivery capacity more efficiently. In this section, we describe preliminary hypothesis we aim to test.

In our analysis we use the following nomenclature: We say that day $d_i$ and $d_j$ are such that $d_i < d_j$ if day $d_i$ occurs earlier than $d_j$. Also, we define the sum $d_j + 3$ to be three days after day $d_j$. Similarly $d_j - 3$ three days earlier than day $d_j$. We use bold letters to denote vectors (e.g. $\mathbf{k}_i$). Element $j$ of $\mathbf{k}_i$ is $k_{i,j}$. Capital letters are reserved for random variables and lower case letters for their realizations. $\mathbb{E}_{\mathbf{D}}$ is the expectation with respect to $\mathbf{D}$.

**The Impact of Lead Time.**

In the context of home delivery operations, several studies have pointed out the benefits of decreasing shipping times and delivery time windows to increase revenues and customer engagement (e.g. [**52**], [**39**], and [**82**]). Waiting time is usually compared to price such that longer lead times are associated with a lower valuation of the delivery service, and it is usually assumed that demand sensitivity to OD-Lead-Time is negative. Hence, it is often argued that if all delivery alternatives have the same price, customers tend to choose delivery alternatives with short OD-Lead-Time. While the argument applies to meal delivery, or non-attended home delivery, where customers do not need to be present at the time of delivery and shippers can leave parcels in lockers, mail rooms or sometimes near the mail box, it is not clear it applies to other delivery services.

We hypothesize that, in AHD, where customers commit to be present on a chosen day, customers might have priorities that conflict with the preference for earlier days. In a different context, we find that patients scheduling non-urgent health care share a similar decision process and trade-offs to customers scheduling AHD. Note that, in health care appointments, patients obviously commit to be present at the time of service delivery. In this context, it is observed that speed of delivery is of limited importance compared to other priorities. Consequently, some studies have found that efforts for increasing speed of service beyond a critical point can be of limited value compared to other needs (e.g., [**109**], [**108**], and [**83**]). Our first hypothesis explores whether similar findings apply to AHD.

***Hypothesis 1:*** *There is a lower bound on OD-Lead-Time such that days offered for delivery too close to the day of purchase are less likely to be chosen. There exist a threshold* $\underline{d}$ *such that if delivery lead time (in days) is shorter than* $\underline{d}$*, the odds that a day is chosen decreases as delivery lead time decreases.*

**The Impact of the Day of Week of the Candidate Day.**

Several studies have explored the impact of the day of week on people's activities in a variety of context; from recreational activities [**16, 17**] to transportation preferences [**141**], they show how people behave differently on weekends than on weekdays. In our context, we hypothesise that convenience for being at home at the time of delivery changes depending on the day of week. We expect Saturday to be more likely to be chosen compared to the rest of days because weekends are more likely to be freer of work commitments. Additionally, conversations with the supply chain team of the furniture company revealed a popular opinion among the delivery teams that Tuesdays are busier than the rest of the days. The reasoning behind this belief is that the company does not deliver on the previous two days, Sunday and Monday, which makes Tuesday the first day of week available for delivery, which might be more attractive for some customer segments. Based on these arguments, we state our second hypothesis.

***Hypothesis 2:*** *Tuesdays and Saturdays are more likely to be chosen when they are offered for delivery compared to Mondays, Wednesdays, Thursdays, and Fridays.*

**The Impact of the Day of Week of Purchase.**

The main channel used by the customers of the furniture company are the retail outlets. The impact of the day of week on customer behavior has been studied in the literature, for instance, when shopping [**47**], or dining out[**41**]. We argue that several factors influence day availability for visiting the stores and making a purchase, such as family commitments, office or working schedules or school's timetable, and that these impact the customer choice of AHD. Many of these factors have a periodic structure that is repeated every week. For instance, working hours typically span from 8:30 a.m. to 4:30 p.m., from Monday to Friday. We hypothesise that the day customers made a purchase provides information about how flexible customers are to choose delivery dates. In particular, we hypothesise that customers that made a purchase on Saturdays or Sundays are more likely to have lower availability (less flexible) and tend to choose later dates and, therefore, they are more likely to reject dates when dates are sequentially offered.

***Hypothesis 3.1:*** *Customers that made a purchase on weekends have, on average, lower availability than customers that made a purchase on weekdays. When a day is offered for AHD, customers that made a purchase on Saturdays or Sundays are more likely to reject it.*

As a finer version of the previous hypothesis, we hypothesize that the day of week customers made a purchase correlates to the customer's likelihood to accept or reject an offered candidate day based on the day of week it occurs. The hypothesis is based on the fact that customer's need to be available to visit the store, as well as to be present at the time of delivery. Customers that are usually unavailable on Tuesdays, are not likely to visit the store on Tuesdays, nor to schedule AHD for Tuesdays.

***Hypothesis 3.2:*** *When a day is offered for AHD, customers that made a purchase on Saturdays or Sundays are more likely to reject a candidate day occurring on a weekday than a candidate day occurring on a weekend.*

Understanding customers' sensibility to fast delivery compared to other priorities such as day of week can allow retailers to balance customer satisfaction and last mile delivery costs.

We find that the day of week that customers visited the store and made a purchase is a strong predictor of the customers' availability. Therefore, the day of week of purchase is a simple and intuitive way to segment customers. Our study provides a method to quantify the effect of day of week on customer availability and its interaction with OD-Lead-Time. Our results also suggest intuitive and easy-to-implement strategies to customize delivery capacity to these customer segments.

## 4.3. Data, Basic Assumptions and Model Description

In Section 4.3.1, we describe the data available by the retail company that we use in our study. In Section 4.3.2 we state the basic assumptions of our model, which we support with empirical results in the appendix. Section 4.3.3 provides more details about the interaction between the customer and the retail company, which we call the *customer-retailer interaction*, and describes our modelling approach.

### 4.3.1 Data Description

To understand how customers chose delivery dates, we analyze a data set from the retail company containing the scheduling transaction from July 2016 to July 2017 in the

Southern US. The area is comprised by nine regions that are supplied by region-specific distribution centers, from where the last mile home delivery routes depart. The total number of transactions in our data set is 30,572. Each transaction entry includes region and store identifiers, open date and time (i.e. day of purchase), date scheduled for delivery and delivery team assignment.

There exist three main holidays sales periods for which the retail company expects a demand surge in all regions simultaneously: Labor Day, Black Friday and Presidents' Day. These expected demand surges impact the subsequent deliveries for the weeks following. Accordingly, the retail company increases the delivery capacity for the periods after these holidays. In Figure 4.1, we show weekly sales and weekly deliveries for the three regions (out of nine included in our data) with the largest sales volume. We excluded the week before and the four weeks after the holidays in our analysis, as both demand and capacity are atypical during these periods. The resulting number of transactions is 19,721.

The focus of our work is to understand how customers choose dates for AHD. Figure 4.2 shows the empirical distribution of the OD-Lead-Time (the days passed between the day of purchase and the AHD) chosen by the customers for the first 28 days. Observe that most customers choose a delivery date within the first four weeks. The distribution for those customers is bimodal and skewed towards the left, showing that dates within the first two weeks tend to be chosen when offered and available. Observe that most customers chose a delivery date within 28 days (four weeks), which represent 91% of our data.

The empirical proportions observed by the retail company are the result of the interaction of two agents: (1) The retail company, who only offers days for delivery that have

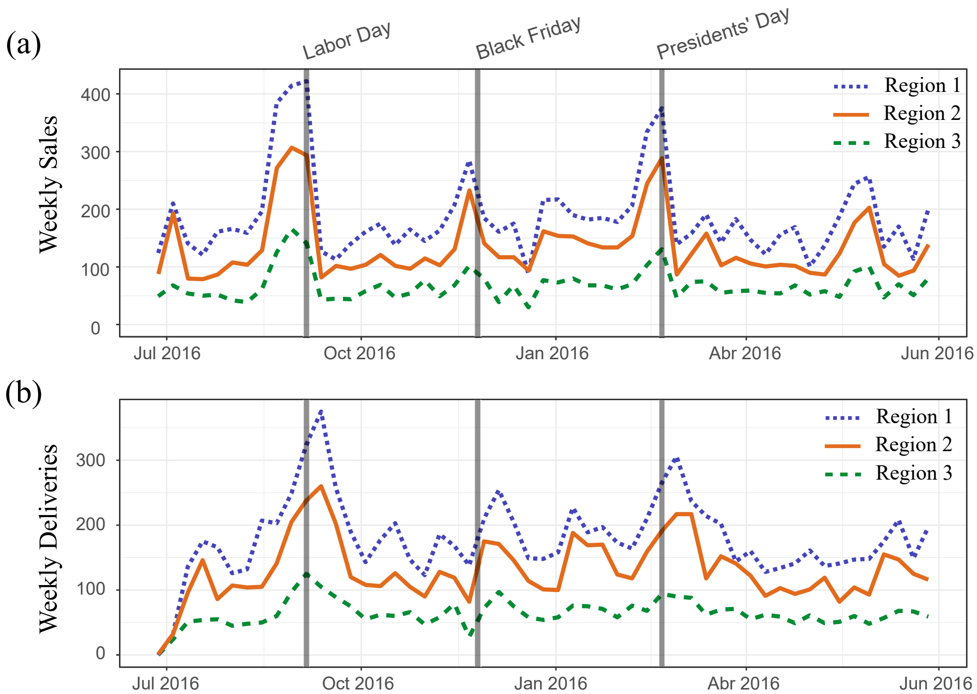


Figure 4.1. (a) Weekly sales for three regions of the retail company. (b) Weekly deliveries for three regions of the retail company.

remaining delivery capacity (those days that are not available due to capacity constraints are not offered to the customer), and (2) the customer, who may choose a delivery date only if such a date is available in the customer's personal schedule. Since the customer must present at the time of delivery, several factors play a role in the customer's decision. Examples of these factors might be speed of delivery, day of week of a candidate date and work or family related commitments.

Since customers are only offered the dates available for the retail company, the empirical proportions observed by the retail company in Figure 4.2-(a) do not fully reveal customers' preferences and availability. Our model aims to provide a framework to infer customer's preferences and availability from observed data, and understand the different factors that lead the customer's choice.

We exclude those customers that scheduled the delivery date beyond 28 days[1]. Our model considers customers that chose as the date of delivery the earliest day available based on the interaction with the retail company, where availability for both parties plays a fundamental role. We observe in our data that beyond the first two weeks, all days have nearly full delivery capacity. It is unlikely that customers that scheduled beyond four weeks chose the earliest available delivery date based on the interaction with the retailer. That would imply that none of the first 28 days were available for the customer. Instead, it is reasonable that the factors driving the choice beyond 28 days are other than availability. For example, as revealed in conversations with the supply chain team of the furniture manufacturer, some customers purchase pieces of furniture for residences under construction or remodelling. For those customers, day of delivery is not driven by customer availability or the interaction with the capacity constraints of the retailer. In that case, a particular date in the calendar is predetermined based on external factors such as construction time. In addition, since days beyond four weeks have full delivery capacity, the proportion of customers choosing late delivery dates beyond the first 28 days can be estimated from the observed empirical proportions, without building a behavioral model.

In contrast, several factors come into a play in the decision of those customers who choose a delivery date within the first two or three weeks. For instance, only a subset of days are available for delivery due to capacity constraints of the retail company. Also, due to the nature of AHD and customer availability, only few of the offered days may be chosen for the customer. In addition, the left-skewed distribution in Figure 4.2 seems to

[1]in Appendix C.1 we show that our analysis can easily incorporate those customers without changing our key insights, which has not been included in the main text to ease the presentation of our results.

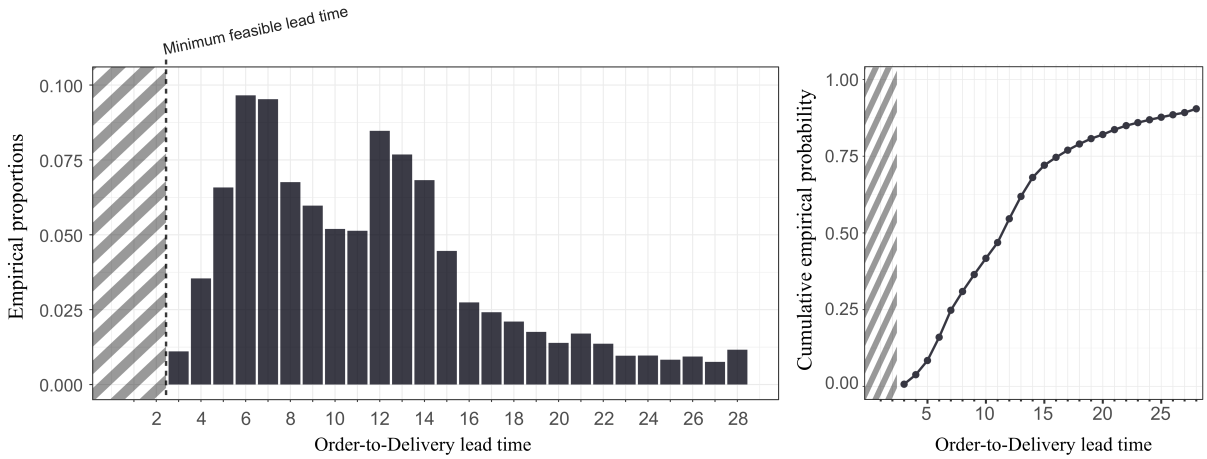


Figure 4.2. Empirical proportions of customer choices for the first 28 days in the schedule.

reveal a preference for earlier dates. Additional factors such as day of week effects may play a role that explains the bimodal shape of the empirical proportions. Understanding these factors is critical to answer the questions on the impact of delivery capacity on customer satisfaction posed by the retail company.

### 4.3.2 Basic Assumptions

Before introducing our model, we first state three key assumptions, which we support using the company's data in appendix C.2, C.3 and C.4.

**Assumption 1: Home Delivery Teams Have Fixed Delivery Capacity.** Our conversations with the supply chain team of the furniture company revealed that home delivery capacity for each day is planned in advance by dedicating delivery teams to regions (a team consists of a van owned by the company, a driver, and technicians). Our data set reports, for each day, the number of delivery teams dedicated to each region. We estimate the delivery capacity of each day by assuming that, for each region, delivery teams have, on average, a fixed delivery capacity. In other words, the average delivery capacity of day $d$ in region $r$ is equal to the per-team delivery capacity, $k_r$, multiplied by the number of delivery teams dedicated to region $r$. Per-team delivery capacity, $k_r$, is a parameter that we estimate using the transactional data (see Appendix C.2) that depends on the region. On average, per-team delivery capacity $k_r$ is about five deliveries per team and seems to vary little across days.

**Assumption 2: The Impact of Lost Sales Is Negligible.** Our model assumes that the immediate impact of delivery lead time and day availability on customer choice of purchasing the product is negligible. In other words, when a customer decides make a purchase and proceeds to closing the sale with the retailer, it is unlikely that the customer withdraws without purchasing (translating into a lost sale) because they are unsatisfied with the home delivery dates offered.

The assumption of negligible lost sales is supported by the nature of the products sold by the furniture company, and the customer experience in-store. The most popular pieces of furniture that require home delivery services are high-quality mattresses and bed frames with retail prices around $1000-$5000, a product that customers do not purchase frequently. Before deciding the home delivery date, customers are assisted by sales representatives who show and describe a variety of products and features. This process takes many minutes, sometimes an hour, and is intended to provide customers with an opportunity to try and have a first-hand experience with the products offered by the company before making a final decision. If a customer decides to purchase a product, they choose features such as color, style, and bedding, and the sales representative starts processing the order into the system. At the final stage, the retailer asks for customer's preferred date to schedule the delivery, a scheduling process that we describe in the next section.

Note that our context is very different from that in fast fashion retailing, flash sales, and similar settings where assortments are constantly changed to increase sales frequency, customers purchase products on the spur of the moment, and lost sales are common (e.g., [**117**] and [**29**]). In contrast, customers of the furniture company carefully test the products and the features that satisfy their needs, because products are expected to be used for years. We assume it is very unlikely that a transaction would translate to a lost sale in the final step because a customer finds unacceptable that the time for delivery is beyond one or two weeks.

To support this assumption, we check whether the number of days with remaining delivery capacity significantly impact sales volume. We find no evidence supporting that lack of delivery capacity increases lost sales. We provide more details in appendix C.3.

**Assumption 3: From among all offered dates, customers choose the earliest available date.** Customers greatly value speed of delivery [**3, 39, 82**], and without availability constraints or pricing differentiation, customers have little incentives for delaying the delivery. Given a customer that chooses between two different days, one day comes earlier than the other, we assume that, if both days are available for the customer, *ceteris paribus*, the customer will more likely choose the earlier date.

In AHD customers need to be present at the time of delivery. Hence, customers must choose an available date from among the offered dates. We assume that, while customers find earlier dates more attractive than later delivery dates, customers choose the date for home delivery based on their availability in AHD. In other words, we assume that customers choose the earliest day offered that is available for them.

Our assumption is also supported by the way dates are typically offered to the customer. Sales representatives usually suggest the earliest with remaining delivery capacity, and ask for the customer's availability. If the first date suggested is not convenient for the customer, future dates are offered sequentially until a date is found to be convenient. Due to this interaction, the scheduled date tends to be the first that is available for the customer among the dates that are available for the retail company.

In Appendix C.4, we show that the data supports this third assumption.

### 4.3.3 The Customer-Retailer Interaction Model

The observed proportions in Figure 4.2 are the result of the interaction between the customer and the retailer that takes place when the customer chooses a date for AHD when finalizing the sale at the company's outlets. We call this interaction the Customer-Retailer Interaction.

Once a sale is finalized at the retail outlets, the Sales Representative (SR) opens a transaction in the system that is filled out with information about the customer (Zipcode, phone number, etc.) and the product. The SR can see the availability of the retail company through the system, i.e.; the days with remaining delivery capacity. With this information, the SR and the customer interact until a delivery date is found to be available for the customer among the proposed dates shown in the system. The interaction typically starts with the SR offering the earliest available date for delivery. If the date proposed is unavailable for the customer, the SR proceeds with the next available day. Such a communication continues until the SR and the customer find an available delivery date. The customer may also start the interaction by proposing a delivery date. The resulting choice is identical in either case.

All dates ($\mathcal{D}_t$): $\{d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9, d_{10}, ...\}$

Retailer Availability Filter ($\mathcal{D}_t^r$): $\{\cancel{d_1}, \cancel{d_2}, d_3, \cancel{d_4}, d_5, \cancel{d_6}, \cancel{d_7}, d_8, d_9, d_{10}, ...\}$

Customer Availability Filter ($\mathcal{D}_t^c$): $\{\cancel{d_1}, d_2, \cancel{d_3}, \cancel{d_4}, \cancel{d_5}, d_6, d_7, d_8, \cancel{d_9}, d_{10}, ...\}$

Final Choice: $\{\cancel{d_1}, \cancel{d_2}, \cancel{d_3}, \cancel{d_4}, \cancel{d_5}, \cancel{d_6}, \cancel{d_7}, \boxed{d_8,} \cancel{d_9}, d_{10}, ...\}$

Figure 4.3. Example of customer choice. Days crossed represent non-available days.

Note that because of the way the SR and the customer interact to find a date for AHD, the scheduled date tends to be the first that is available for the customer among the dates that are available for the retail company. We model the interaction between the SR and the customer as follows, and we use the example in Figure 4.3 to illustrate our model. We let a customer, $c_t$, and the SR start interacting at time $t$ of day $d$ (i.e.; customer makes a purchase at time $t$). We define the day $d_j$ as the day that is $j$ days after the purchase date ($d_j = d + j$). We denote the set of all candidate days at time $t$ as $\mathcal{D}_t = \{d_1, d_2, \ldots, d_{28}\}$, shown in the first row in Figure 4.3. The SR sees in the system the subset of available dates of $\mathcal{D}_t$, which we denote as *Retailer Availability Filter* $\mathcal{D}_t^r \subseteq \mathcal{D}_t$, shown in the second row. With $\mathcal{D}_t^r$, the customer and the SR interact until they find the first day in $\mathcal{D}_t^r$ that is available for the customer. More specifically, from all days $\mathcal{D}_t$, only a subset are available for the customer, which we denote as *Customer Availability Filter* $\mathcal{D}_t^c \subseteq \mathcal{D}_t$, shown in the third row. The final choice after the customer-retailer interaction, $d_t^*$, is the first day in the intersection of the two filters $\mathcal{D}_t^r \cap \mathcal{D}_t^c$, which are the dates that are available for both the retail company and the customer ($d_t^* = \min_{d_j \in \mathcal{D}_t^r \cap \mathcal{D}_t^c} j$), corresponding to the eighth day $d_8$ in our example.

With these two filters, the customer-retailer communication during the interaction could have been as follows: The SR offers the first day available for the retail company, $d_3 (\in \mathcal{D}_t^r)$, shown in the system. However, such a date is not available for the customer ($d_3 \notin \mathcal{D}_t^c$) and, therefore, the customer rejects the offer. The next date offered to the customer is $d_5$, which, again, is not available for the customer ($d_5 \notin \mathcal{D}_t^c$) and is rejected. In a third attempt, the SR offers day $d_8$. Since $d_8$ is available for the customer ($d_8 \in \mathcal{D}_t^c$), the date is accepted for delivery. The final choice would have been the same, had the customer started the interaction asking for the retailer availability of the first day available $d_2 \in \mathcal{D}_t^c$.

Unfortunately, the exact communication during the interaction is not registered in the system and, therefore, it is not visible in the transactional data from the retail company. Transactional data only reveals the final choice of the customer $d_t^*$. The retailer availability $\mathcal{D}_t^r$, used in combination with the final choice, $d_t^*$, only reveals that the previous days available for the retail company, $d_3$ and $d_5$ in the example of Figure 4.3, were not available for the customer, i.e.; $d_3 \notin \mathcal{D}_t^r$ and $d_5 \notin \mathcal{D}_t^r$. For example, had $d_3$ been offered, the customer would have chosen it for delivery.

Therefore, from the transactional data and the home delivery capacity, the Retailer Availability filter in any transaction, $\mathcal{D}_t^r$, is fully observed. In contrast, only a small piece of information about the Customer Availability Filter $\mathcal{D}_t^c$ is revealed from the customer choice. In the next section we show that the retail company can combine and use the information provided by each individual transaction to infer on the customer availability filter and better understand the customers' choices empirically observed.

## 4.4. Research Methods

In this section we first show how the data from the company can be used to infer on the customer availability filter, we then contextualize our estimation framework, and describe how results and parameters can be interpreted from our model.

### 4.4.1 Estimation of Customer Availability Filter

In AHD, customers must be present at the time of delivery and, when considering a candidate delivery date, customers evaluate if such a day is available in their personal schedule. The subset of days for which the customer can be at home determines the Customer Availability Filter $\mathcal{D}_t^c$. Unfortunately, this subset of days cannot be fully observed from our transaction data. In this section we propose a model to estimate the probability that a candidate day is available for the customer, i.e.; the probability that a candidate day belongs to the Customer Availability Filter $\mathcal{D}_t^c$.

For a candidate day $d_j$ in the set of all days $\mathcal{D}_t$, the probability that such a day is available for a customer depends on several factors. For instance, it may depend on customer features, such as employment characteristics, family structure or type of house. It may also depend on features of the candidate day $d_j$, such as OD-Lead-Time (i.e., difference in days between date of purchase and $d_j$, which is the integer $j$) or candidate day of week. Availability of a candidate day can also depend on interactions between the two sets of features. For instance, customers with a rigid work schedule (customer feature) may have lower availability on Tuesdays (candidate day feature) than those customers with more flexible work schedules.

We model the probability of availability as follows. We let $c_t$ be the customer that makes a purchase at time $t$ with a SR. The probability that candidate day $d_j$ belongs to the Customer Availability Filter $\mathcal{D}_t^c$ of customer $c_t$ is $P[d_j \in \mathcal{D}_t^c | d_j, c_t]$. In Equation (4.1) we model this probability with a function, $\pi(\bullet)$, that depends on features of the candidate date and the customer, $\mathbf{X}_{j,t}$, and a set of parameters $\boldsymbol{\beta}$. In our study, we use a logistic function that depends on the product $\boldsymbol{\beta}'\mathbf{X}_{j,t}$.

$$P[d_j \in \mathcal{D}_t^c | d_j, c_t] \sim \pi(\mathbf{X}_{j,t} | \boldsymbol{\beta}) = \frac{\exp(\boldsymbol{\beta}'\mathbf{X}_{j,t})}{1+\exp(\boldsymbol{\beta}'\mathbf{X}_{j,t})} \tag{4.1}$$

We aim to estimate the parameters $\boldsymbol{\beta}$ that best explain the data observed from the retail company. To have a finite set of alternatives, the last day, $d_{28}$ includes customers that scheduled a delivery on the 28th day and beyond. This is equivalent to including an "outside option" in the choice literature[2].

A candidate day $d_j$ needs to meet three conditions to become the final choice of the customer after the interaction with the retailer:

- Available for the customer ($d_j \in \mathcal{D}_t^c$).
- Available for the retail company ($d_j \in \mathcal{D}_t^r$).
- Offered by the SR: Assuming that from all available dates, the earliest is chosen (third assumption in Section 4.3.2), a day is offered only if none of the previous

[2]Alternatively we could include an outside option as Day 29 representing the option that no day has been chosen within the first 4 weeks. The analysis and results obtained would be equivalent.

dates are chosen. In other words, the candidate day $d_j$ may be offered only if the previous dates are not available for delivery either for the customer, the retailer or both ($d_i \notin \mathcal{D}_t^r \cap \mathcal{D}_t^c$ for all $i < j$).

The first two conditions imply that the candidate day $d_j$ belongs to the intersection of $\mathcal{D}_t^c$ and $\mathcal{D}_t^r$. The third condition implies that, for day $d_j$ to be considered, all the previous candidate days must not belong to the intersection of $\mathcal{D}_t^c$ and $\mathcal{D}_t^r$. Therefore, the probability that candidate day is chosen depends on the availability of the previous days (Assumption 3).

Assuming that the probability of availability of a candidate day $d_j$ is independent of the previous dates (condition on the customer $c_j$), the probability that a candidate day is chosen can be easily computed. We remark that, even though the probability of availability (condition on the customer $c_j$) is independent of the previous days, the probability that a candidate is chosen still depends on the availability of previous dates. By denoting $D^*$ the final choice of the customer, in Equation (4.2), we show that the probability of the candidate day $d_j$ being chosen, $P[D^* = d_j | X_{j,t}]$, is a sequence of conditional probabilities.

$$\begin{aligned} &P[D^* = d_j | X_{j,t}] = \pi(\mathbf{X}_{j,t} | \boldsymbol{\beta}) P[D^* \neq d_{j-1} | X_{j-1,t}] && \text{if } d_j \in \mathcal{D}_t^r, \\ &P[D^* = d_j | X_{j,t}] = 0 && \text{if } d_j \notin \mathcal{D}_t^r. \end{aligned} \tag{4.2}$$

With initial conditions,

$$\begin{aligned} &P[D^* = d_1 | X_{1,t}] = \pi(\mathbf{X}_{1,t} | \boldsymbol{\beta}) && \text{if } d_1 \in \mathcal{D}_t^r, \\ &P[D^* = d_1 | X_{1,t}] = 0 && \text{if } d_1 \notin \mathcal{D}_t^r. \end{aligned}$$

In Equation (4.2), the probability that day $d_j$ is chosen is the product of the probability that the day is available for the customer and the probability that the previous dates are not chosen. By construction, $\sum_{j=1}^{28} P[D^* = d_j | X_{j,t}] = 1$ for any vector of parameters $\boldsymbol{\beta}$. Hence, our model provides well defined choice probabilities without adding additional constraints often required in non-parametric estimation approaches (e.g.; [**131**] and [**99**]).

To estimate the model parameters $\boldsymbol{\beta}$, we maximize the likelihood function of the transactional data from the retail company. Let $N_r$ be the total number of transactions of region $r \in \{1, \ldots, 9\}$, and $d_{o,t}$ be the observed customer's choice of transaction $t \in \{1, \ldots, N_r\}$. We write in Equation (4.3) the expression of the likelihood function.

$$\mathcal{L}(\boldsymbol{\beta}) = \prod_{r=1}^{9} \prod_{t=1}^{N_r} P[D^* = d_{o,t} | X_{j,t}]. \tag{4.3}$$

It can be shown that the function $\mathcal{L}(\boldsymbol{\beta})$ is concave in the parameters $\boldsymbol{\beta}$ and, therefore, it is easily maximized using standard methods and packages. The maximizer $\boldsymbol{\beta}^*$ provides asymptotically optimal estimates of the true parameters [**68**]. [**53**], [**80**] and [**63**] show how robust and cluster-robust standard errors can be obtained for maximum likelihood estimators with this structure.

The following section describes the set of variables $\mathbf{X}_{j,t}$ that are included in the availability model $P[D^* = d_j | X_{j,t}]$.

### 4.4.2 Customer and Candidate Day Features

Our data contains 17,751 transactions, which we index by $t$. The probability model of a candidate day $d_j$ being available for customer $c_t$, $\pi(\mathbf{X}_{j,t}|\boldsymbol{\beta})$, depends on features of the customer and the candidate day, $\mathbf{X}_{j,t}$, which we briefly describe below.

**Customer Features.** We summarize the customer features in Table 4.1 and we describe each feature below.

- **Region**: There are a total of 9 regions. For each region $r$, we incorporate a dummy variable $Reg_{r,t}$ , such that $Reg_{r,t} = 1$ if transaction $t$ was opened in region $r$, and 0 otherwise.
- **Seasonality**: Our data spans from July 2016 to July 2017, with sales periods removed. We split the season in 22 biweek windows. We incorporate a dummy variable $Season_{s,t}$ for each biweek window $s$, such that $Season_{s,t} = 1$ if transaction $t$ was opened in the window $s$, and 0 otherwise.
- **Time of the day**: We split the day in three time windows indexed with $q$: (1) Earlier than 12 p.m. (2) 12 p.m.-4 p.m. (3) later than 4 p.m. We incorporate a dummy variable $Time_{q,t}$ for each time window $q$, such that $Time_{q,t} = 1$ if transaction $t$ was opened during time window $q$, and 0 otherwise.
- **Day of Week of Purchase**: We incorporate a dummy variable $DoWP_{w,t}$ for each day of week $w$, such that $DoWP_{w,t} = 1$ if transaction $t$ was opened during time window $w$, and 0 otherwise ($w = \{Mon, Tue, \ldots, Sun\}$).

**Candidate Day Features.** We focus our attention on those customers that scheduled the delivery within 28 days from the day of purchase. Each candidate day $d_1, \ldots, d_{28}$ has features that impact customer availability. We describe the features below.

| Variable Name | Num. of Variables | Av # of obs. | sd. # of obs | Min # of obs. | Max # of obs |
|---|---|---|---|---|---|
| $Season_{s,t}$ | 21 | 845 | 308 | 338 | 1613 |
| $Time_{q,t}$ | 3 | 5917 | 883 | 4898 | 6433 |
| $Reg_{r,t}$ | 9 | 1972 | 1902 | 571 | 5943 |
| $DoWP_{w,t}$ | 7 | 2536 | 820 | 1983 | 4254 |

Table 4.1. Summary of customer features variables.

- **OD-Lead Time**: To capture the effect of OD-Lead Time of the candidate days (i.e.; the days passed since the time of purchase), we incorporate the logarithm of OD-Lead Time as a continuous variable. For day $d_j$, the logarithm of OD-Lead Time is $logODLT_j = \log(j)$.
- **Candidate Day of Week**: We incorporate a dummy variable $DoWCD_{v,j}$ for each day of week $v$, such that $DoWCD_{v,j} = 1$ if the candidate day of week $d_j$ is $v$, and 0 otherwise. Sunday and Monday are not included in the set of variables because the delivery teams of the retail company did not deliver on Sundays or Mondays ($v = \{Tue, Wed, Thu, Fri, Sat\}$).

For each set of binary variables, we drop the first index to indicate its vector form, e.g.; $\mathbf{Season}_t = (Season_{1,t}, Season_{2,t}, \dots)$.

In our analysis, we use the marginal effects of the customer and candidate day features to gain insight onto the drivers of customer choice. As we describe in the next section, marginal effects cannot be directly evaluated from $\boldsymbol{\beta}^*$ because we model the decision as a sequence of conditional probabilities (see Eq. 4.2). This sequential model belongs to the family of sequential logistic models, which have been largely used in the literature. In the next section we discuss how this model can be used to interpret the impact of customer and candidate day features, $X_{j,t}$.

### 4.4.3 Logistic Sequential Models: Context and Interpretation

Sequential logistic models have been used in the literature to model situations or processes in which agents make decisions of continuing or leaving the system in an ordered sequence. In our context, the customer sequentially accepts or rejects the candidate days offered by the SR. If a candidate day offered is available for the customer, the day is accepted for home delivery and the customer "leaves" the system; otherwise, the customer-retailer interaction continues.

Similar models have been widely used in the education literature to model how students progress across school grades, and to measure the impact of student characteristics (i.e., economic or family background) on the decision of continuing to the next grade or leaving the school (e.g., [**89**], [**90**], [**137**], and [**107**]). In this context, after completing an academic grade, students may decide to stop their education ("leave the system") or to continue to the next grade. This model approach for education attainment became popular and the literature termed it the "Mare Model" after [**89**]. In a different context, more recently, [**116**] uses similar models to estimate innovation success. The authors measure the impact of personal treats on the success rate in three ordered stages leading to product innovation: idea generation, prototyping and diffusion. In this context, individuals engaged in new product development at a given stage may decide to stop the process and drop the product idea, or continue to the next stage.

[**27**] shows that due to heterogeneity, selection bias and omitted variables, the estimated parameters ($\boldsymbol{\beta}$ in our model) of sequential logit models cannot be directly interpreted as the marginal effects of the features. More specifically in our model, the magnitude of the parameters is linked with the choice heterogeneity of the customers

[**124**]. Therefore, the comparison of parameters between groups of customers (e.g. customers that purchased the product on Saturday vs. on Tuesday) is challenging because different groups may have different heterogeneity. In addition, the impact of parameters for late candidate days is particularly challenging to interpret because only customers with very low availability are observed, introducing selection bias. Therefore, only a subset of customers is observed for those dates. Finally, [**27**] and [**24**] show that omitted variables can also bias the magnitude of the parameters.

To overcome these challenges, [**27**] shows that marginal effects can be reliably estimated by considering changes in probability terms, also known as *Partial Effects* (see also [**63**], [**95**] and [**96**]). Following this approach, we measure the impact of customer and candidate features on customer availability as follows: Let $x_{j,t,k}$ be a feature $k$ of the feature vector $\hat{\mathbf{x}}_{j,t}$. If $x_{j,t,k}$ is a continuous feature, we evaluate the marginal effects of $x_{j,t,i}$ with the partial effect calculated as $\frac{\partial \pi(\hat{\mathbf{x}}_{j,t}|\boldsymbol{\beta})}{\partial x_{j,t,k}}$. If $x_{j,t,k}$ is a binary feature, we evaluate the marginal effects of $x_{j,t,i}$ with the partial effect calculated as $\pi(\hat{\mathbf{x}}_{j,t}|x_{j,t,k}=1)-\pi(\hat{\mathbf{x}}_{j,t}|x_{j,t,k}=0)$. [**63**] shows how cluster robust standard errors for the partial effects can be calculated from the previous expressions. The partial effects depend on the point they are evaluated, $\hat{\mathbf{x}}_{j,t}$, therefore, each time we show partial effects, we specify the point, $\hat{\mathbf{x}}_{j,t}$, on which the partial effects are evaluated.

There are several factors that may influence customers' availability for home delivery. Partial effects allow us to evaluate how probability of availability changes for different factors such as day of week of a candidate day or OD-Lead-Time. We can also evaluate how probability of availability changes for different customer groups, such as customers that made a purchase on weekdays versus those that made a purchase on weekends.

In Section 4.5, we follow the approach presented in this section to estimate the impact of these factors on the customer availability filter $\mathcal{D}_t^c$, which, along with the retailer availability filter $\mathcal{D}_t^r$, leads to the final choice.

## 4.5. Findings

In this section we describe the results of several specifications of customer availability. The insights and implications of our findings are discussed in Section 4.6

### 4.5.1 The Impact of OD-Lead-Time and Day of Week

We first explore the effect of OD-Lead-Time of the candidate day and the day of week (DoW). To test the significance of our variables of interest, OD-Lead-Time and DoW, in increasing the maximum likelihood, we use a base model, *Model-0*, which only incorporates control variables for time, seasonality and region (the first three rows of Table 4.1). Therefore, for *Model-0*, the term $\boldsymbol{\beta}' X_{j,t}$ can be expressed as in Equation (4.4).

$$\text{Model 0:}\quad \boldsymbol{\beta}' X_{j,t} = \boldsymbol{\beta}'_S \boldsymbol{Season}_t + \boldsymbol{\beta}'_{Time} \boldsymbol{Time}_t + \boldsymbol{\beta}'_{Reg} \boldsymbol{Reg}_t \equiv \boldsymbol{\beta}'_c \boldsymbol{Control}_t. \tag{4.4}$$

It is well known that waiting time can have an impact on customer choice (e.g., [**83**] and [**64**]). To test hypothesis 1 we include in our model the impact of time passed between the day of purchase and a candidate day, the OD-Lead-Time. Incorporating this variable allows availability of early candidate days to be different than late candidate days. Also, to capture the day of week heterogeneity and to test hypothesis 2 and 3.1, we incorporate the effect of the candidate DoW and DoW of purchase in the model. Model LD-DoW in Equation (4.5) incorporates the control variables of Model 0 along with the effect of OD-Lead-Time as well as DoW.

$$
\begin{aligned}
\text{Model LD-DoW:} \quad \boldsymbol{\beta}' X_{j,t} = & \beta_{LT} logODLT_j + \boldsymbol{\beta}'_{DoWCD} \boldsymbol{DoWCD}_j + \\
& \boldsymbol{\beta}'_{DoWP} \boldsymbol{DoWP}_t + \boldsymbol{\beta}'_c \boldsymbol{Control}_t,
\end{aligned} \tag{4.5}
$$

where, $\boldsymbol{\beta}'_C \boldsymbol{Control}_{j,t}$ is the set of controls of Model 0 in Equation (4.4). We show a summary of Model 0 and Model LD-DoW in Table 4.2.

We analyze the marginal effect of the variables of our model in probability terms following [**27**] and [**63**]. To that aim, we evaluate the change of the variables of interest (OD-Lead-Time and DoW) while averaging the impact of the rest. For instance, if $\overline{r}$ is the total number of region-specific binary variables ($Reg_{r,t}$) and $\boldsymbol{\beta}_{Reg,r}$ is the element $r$ of the vector of parameters $\boldsymbol{\beta}_{Reg}$, the averaged effect of region is calculated as $\frac{1}{\overline{r}} \sum_{r=1}^{\overline{r}} \boldsymbol{\beta}_{Reg,r}$.

In Figure 4.4-(a), we show how probability of availability changes in the first 14 days (two weeks) with OD-Lead-Time, when the rest of the effects are averaged. Observe that probability of availability increases with OD-Lead-Time from, approximately, 0.06($\pm$0.02) to 0.23($\pm$0.01) after two weeks, an increase of about 450%. In Figure 4.4-(b), we show the partial effects of the OD-Lead-Time, which can be interpreted as the increment of probability for each unit of OD-Lead-Time, i.e.; the derivative of the curve in Figure 4.4-(a). The effect of time is particularly strong in early dates, where probability of availability increases, on average, by 0.018($\pm$0.001) for each extra day of OD-Lead-Time. For later dates, OD-Lead-Time is lower in magnitude but still significant 0.013($\pm$0.004).

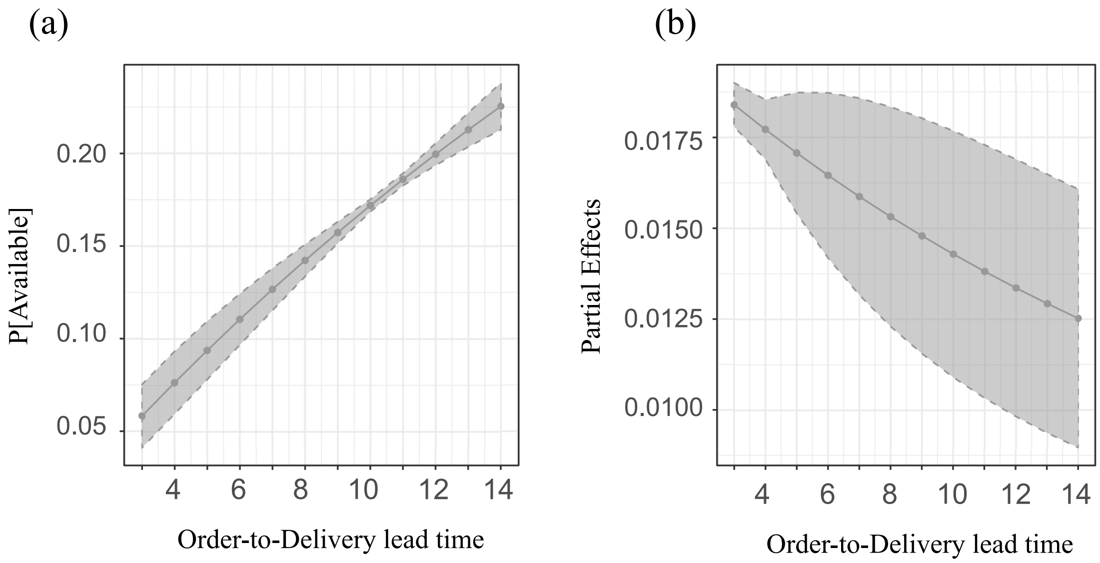


Figure 4.4. Example of customer choice. Days crossed represent non-available days. Shaded region shows cluster robust standard errors, clustered by group, candidate DoW, and purchase DoW.

We evaluate the effect of DoW by comparing the results of Model LD-DoW with Model DoW and Model LD, in Table 4.2. Model DoW is identical to Model 0 but includes the effect of DoW of purchase and candidate day. Similarly, Model LD is identical to Model 0 but includes the effect of OD-Lead-Time. Observe that, while significant, including the set of variables of DoW of purchase and candidate day has small impact on the likelihood with respect to Model 0. Most of the improvement of Model LD-DoW can be attributed to the single variable that captures the effect of OD-Lead-Time. This is confirmed by the increase in likelihood obtained in Model LD.

We further analyse the effect of DoW in Model LD-DoW by means of the partial effects. Figure 4.5-(a) and (b) show how probability of availability changes when we single out the effect by candidate DoW. Observe that the partial effects are overlapped, indicating that there are small differences in availability by DoW.

| | Var. name | Model 0 | Model DoW | Model LD | Model LD-DoW |
|---|---|---|---|---|---|
| **Intercept** | | ✓ | ✓ | ✓ | ✓ |
| **Season** | $\boldsymbol{Season_t}$ | ✓ | ✓ | ✓ | ✓ |
| **Region** | $\boldsymbol{Reg_t}$ | ✓ | ✓ | ✓ | ✓ |
| **Time Visit** | $\boldsymbol{Time_t}$ | ✓ | ✓ | ✓ | ✓ |
| **DoW purchase** | $\boldsymbol{DoWP_t}$ | | ✓* | | (Table 4.3)* |
| **DoW Candidate day** | $\boldsymbol{DoWCD_j}$ | | ✓* | | (Table 4.3)* |
| **OD-Lead-Time** | $logODLT_j$ | | | 0.016 (0.003)* | 0.016 (0.003)* |
| **N obs.** | | 17751 | 17751 | 17751 | 17751 |
| **N pars.** | | 31 | 41[(a)] | 32 | 42[(a)] |
| **LogLikelihood** | | -48182 | -48119 | -45867 | -45829 |
| **Pseudo R-Squared** | | 0.004 | 0.005 | 0.052 | 0.053 |
| $\boldsymbol{\chi}^{2\text{(b)}}$ | | 4984.73 | 4861.13 | 1611.85 | 1557.41 |

* The incorporation of the variables provides significant reduction ($p < 10^{-3}$) of the likelihood based on the likelihood ratio test.

(a) It includes 10 DoW-related variables: 6 $DoWP_t$ (7 - 1 due to the intercept) and 4 $DoWCD_j$ (5 - 1 due to the intercept).

(b) In reference to Figure 4.6.

Table 4.2. Comparison and results of Model 0, Model DoW, Model LD and Model LD-DoW.

In Figure 4.5-(c) we show the averaged partial effects of different candidate DoW when OD-Lead-Time is 7. Table 4.3 show the results in terms of average probability and the partial effects. A positive (negative) partial effect indicates an increase in the average availability of a candidate day. The impact of DoW of the candidate day is small, close to zero in most cases. In Figure 4.5-(d), similar results are obtained when we single out the effects by day of purchase.

To better understand the contribution of the variables to explaining customer behavior, we simulate customer choice according to our model, using the specifications in Table 4.2. In Figure 4.6-(a) and (b), we compare the results obtained from our model with the empirical proportions observed in our data. The last row of Table 4.2 reports the $\boldsymbol{\chi}^2$ statistic. The models that incorporate the impact of OD-Lead-Time provide more

| | DoW | Probability | PE[a] | t-stat PE[a] |
|---|---|---|---|---|
| DoW visit | Monday | 0.129 (0.013) | 0.008 (0.011) | 0.712 |
| | Tuesday | 0.128 (0.014) | 0.006 (0.012) | 0.544 |
| | Wednesday | 0.129 (0.015) | 0.008 (0.012) | 0.623 |
| | Thursday | 0.123 (0.012) | 0.002 (0.009) | 0.225 |
| | Friday | 0.125 (0.012) | - | - |
| | Saturday | 0.117 (0.014) | -0.004 (0.011) | -0.377 |
| | Sunday | 0.130 (0.014) | 0.008 (0.012) | 0.716 |
| DoW alternative | Tuesday | 0.124 (0.015) | -0.010 (0.012) | -0.841 |
| | Wednesday | 0.115 (0.011) | -0.019 (0.011) | -1.700 |
| | Thursday | 0.122 (0.011) | -0.012 (0.011) | -1.096 |
| | Friday | 0.132 (0.012) | -0.002 (0.011) | -0.202 |
| | Saturday | 0.125 (0.012) | - | - |

[a] Note that since DoW variables are binary, partial effects are always expressed as a change relative to an arbitrary reference case, which corresponds to the variables with missing PE. Without loss of generality, they are Saturday (for the candidate DoW variables) and Friday (for the DoW of purchase variables).

Table 4.3. Partial effects by candidate DoW and DoW of purchase of Model LD-DoW. OD-LD is taken at the value of 7 and the rest of variables are averaged. Parentheses show cluster robust standard errors.

accurate predictions of the choice proportions of each candidate day, compared to the empirical data.

The empirical proportions in Figure 4.6 display a bimodal shape that is not captured by our models. The first mode occurs on days 6 and 7 after purchase, and the second mode occurs on days 12 and 13. This seems to indicate that there exists a periodicity in our data, driven by the DoW, that is not captured by the models presented in Table 4.2.

The models in Table 4.2 assume that the impact of DoW is homogeneous. That is, $DoWCD_j$, which captures heterogeneity in availability of the candidate DoW, has the same effect for all customers. Similarly, $DoWP_t$, which captures heterogeneity in availability by DoW of purchase, has the same effect on the availability of all candidate

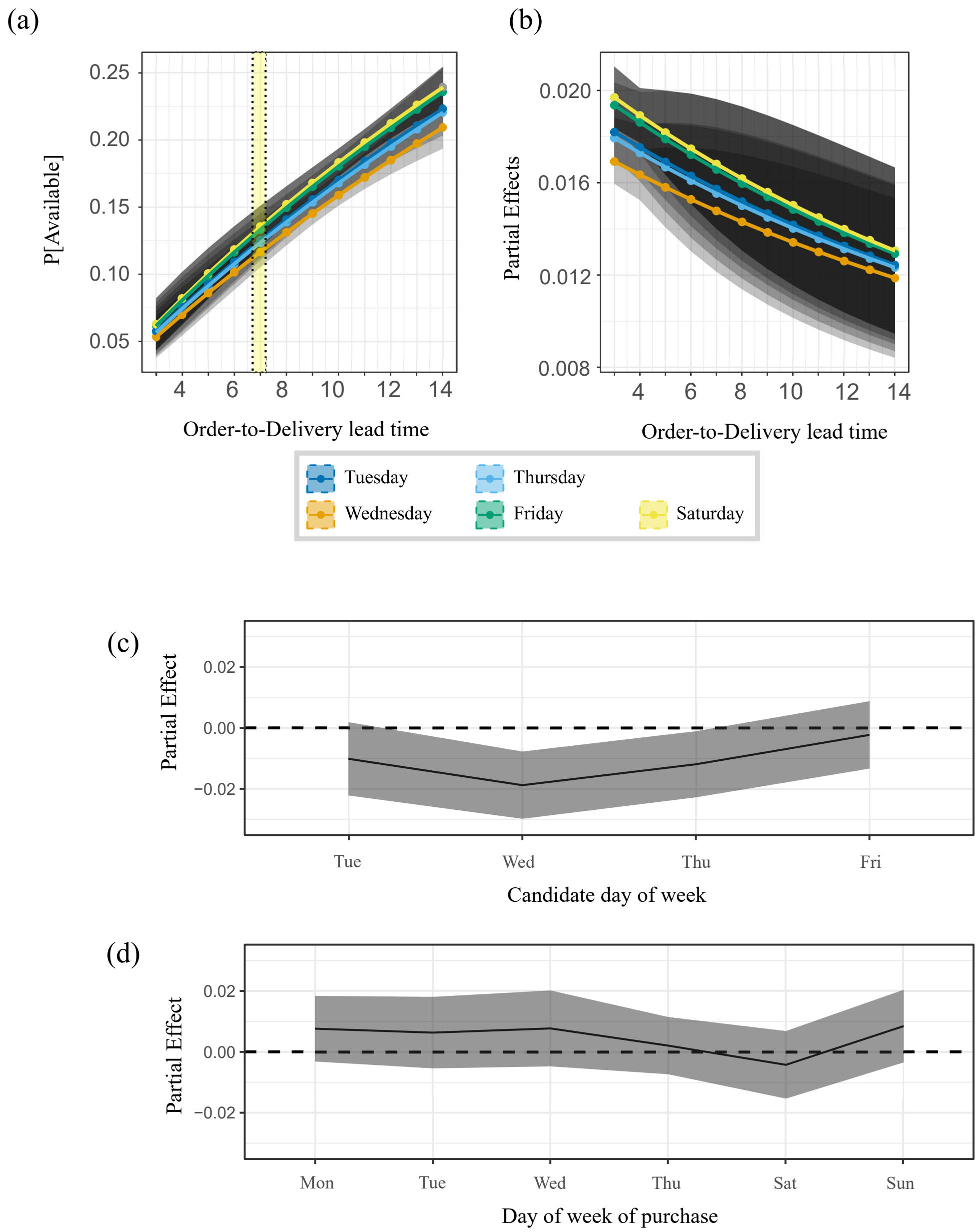


Figure 4.5. (a) The impact of candidate DoW and OD-Lead-Time. Shadows and errorbars show the standard errors, (b) partial effects of the candidate DoW relative to Saturday. (c) partial effects of the DoW of purchase relative to Friday.

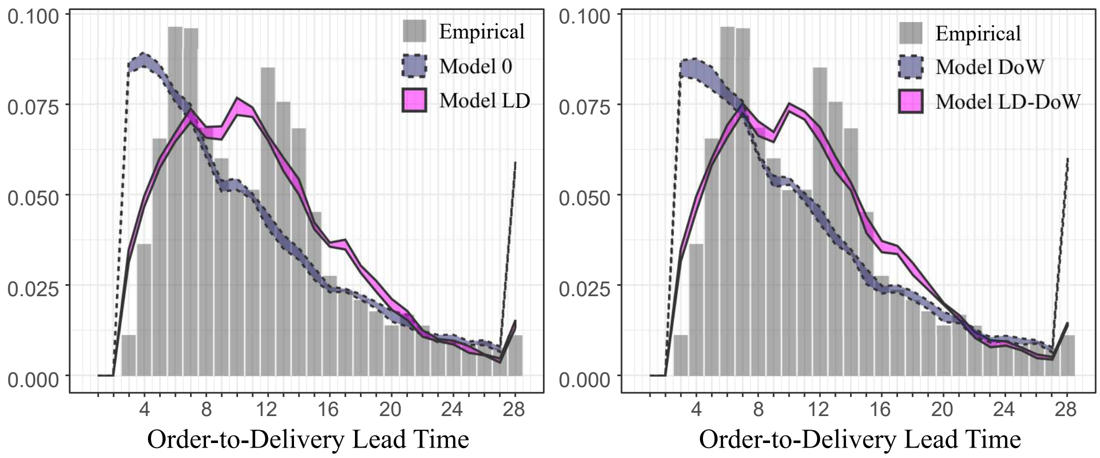


Figure 4.6. Predicted proportions. Results report the mean ± standard deviation of the proportions obtained of 5 simulations.

days. This limitation explains the small contribution of DoW variables to predicting the empirical proportions.

In the next step of our analysis we incorporate customer-specific candidate day availability variables, which allows different customer segments to have different candidate day availability.

### 4.5.2 The Impact of Crossed Day of Week Terms

To explore hypothesis 3.2 and capture the heterogeneity due to day-specific availability by DoW of purchase, we introduce a set of interaction (or "crossed") variables of DoW, $\boldsymbol{XDoW}_{j,t}$. For each pair $w, v$ of DoW of purchase and candidate day DoW, we introduce a binary variable $XDoW_{w,v,j,t}$ such that, $XDoW_{w,v,j,t} = 1$ if customer $j$ made a purchase on the DoW $w$ and if candidate day $d_t$ is on the DoW $v$, and 0 otherwise. The total number of variables is 35.

| | Var. name | Model XDoW | Model LD-XDoW |
|---|---|---|---|
| **Intercept** | | ✓ | ✓ |
| **Season** | $\boldsymbol{Season}_t$ | ✓ | ✓ |
| **Region** | $\boldsymbol{Reg}_t$ | ✓ | ✓ |
| **Time Visit** | $\boldsymbol{Time}_t$ | ✓ | ✓ |
| **Cross-DoW** | $\boldsymbol{XDoW}_{j,t}$ | ✓* | (Table 4.5) * |
| **OD-Lead-Time** | $logODLT_j$ | | 0.015 (0.001)* |
| **N obs.** | | 17751 | 17751 |
| **N pars.** | | 65 | 66 |
| **LogLikelihood** | | -46996 | -45122 |
| **Pseudo R-Squared** | | 0.025 | 0.063 |
| $\boldsymbol{\chi}^{2(a)}$ | | 3444.25 | 492.41 |

* The incorporation of the variables provides significant reduction ($p < 10^{-3}$) of the likelihood based on the likelihood ratio test.

(a) In reference to Figure 4.9.

Table 4.4. Comparison and results of Model XDoW and Model LD-XDoW. Partial effects shown in table (OD-LD is taken at the value of 7 and the rest of variables are averaged). Parentheses show cluster robust standard errors.

In Table 4.4 we show the results of two models that incorporate the crossed variables. Model XDoW incorporates the crossed variables $\boldsymbol{XDoW}_{j,t}$ but it does not include the effect of OD-Lead-Time. Model LD-XDoW incorporates both the crossed variables and OD-Lead-Time.

Compared to Model DoW (Table 4.2), Model XDoW significatively increases the likelihood by capturing the heterogeneity of day-specific availability by DoW of purchase. In Model LD-XDoW, OD-Lead-Time has a significant impact compared to Model XDoW.

We analyze the average impact of OD-Lead-Time on availability of Model LD-XDoW in Figure 4.7-(a) and -(b), which shows the average probability of availability and the partial effects. The impact of OD-Lead-Time of Model LD-XDoW is similar to Model LD-DoW: The effect of OD-Lead-Time is strong, particularly in early dates, in which averaged probability of availability increases from 0.06 (± 0.01) to 0.22 (± 0.22) in two weeks.

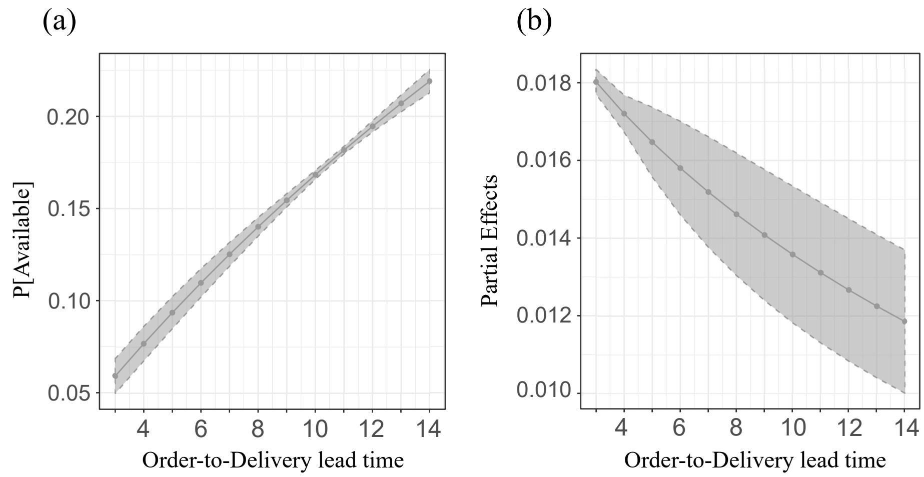


Figure 4.7. The impact of OD-Lead-Time in Model LD-XDoW. Shadow shows cluster robust standard errors.

We detail the effect of the crossed DoW in Figure 4.8 and Table 4.5. Figure 4.8-(b) shows the results for customers that visit the store to make a purchase on Tuesdays. Probability of availability of a candidate day changes depending on the DoW. Customers that made a purchase on Tuesdays have particularly high availability on Tuesdays and Wednesdays (partial effect is high). In contrast, Figure 4.8-(f) shows that, for customers that made a purchase on Saturdays, availability for Fridays and Saturdays is particularly high. Customers that purchased the products in different DoW have different DoW availability.

In Figure 4.9, we show how the day of week interaction variables contribute to explaining the empirical proportions observed in our data. Observe in Model XDoW that the crossed DoW partially captures the two modes of the distribution occurring in OD-Lead-Time 6-7 and 13-14. However, the predicted proportions of people choosing earlier dates

| DoW Visit | Candidate DoW | Probability | PE[a] | t-stat PE[a] |
|---|---|---|---|---|
| Monday | Tuesday | 0.166 (0.013) | 0.079 (0.015) | 5.24 |
| | Wednesday | 0.126 (0.010) | 0.039 (0.013) | 2.95 |
| | Thursday | 0.084 (0.005) | -0.004 (0.014) | -0.26 |
| | Friday | 0.110 (0.004) | 0.022 (0.014) | 1.60 |
| | Saturday | 0.181 (0.012) | 0.094 (0.017) | 5.66 |
| Tuesday | Tuesday | 0.143 (0.012) | 0.056 (0.015) | 3.80 |
| | Wednesday | 0.115 (0.012) | 0.027 (0.015) | 1.84 |
| | Thursday | 0.096 (0.006) | 0.008 (0.015) | 0.54 |
| | Friday | 0.116 (0.008) | 0.028 (0.015) | 1.94 |
| | Saturday | 0.181 (0.015) | 0.094 (0.017) | 5.65 |
| Wednesday | Tuesday | 0.132 (0.010) | 0.044 (0.014) | 3.13 |
| | Wednesday | 0.115 (0.009) | 0.027 (0.013) | 2.11 |
| | Thursday | 0.081 (0.004) | -0.006 (0.014) | -0.45 |
| | Friday | 0.183 (0.013) | 0.096 (0.016) | 6.07 |
| | Saturday | 0.143 (0.017) | 0.055 (0.019) | 2.84 |
| Thurday | Tuesday | 0.124 (0.013) | 0.037 (0.016) | 2.28 |
| | Wednesday | 0.117 (0.009) | 0.030 (0.013) | 2.28 |
| | Thursday | 0.118 (0.008) | 0.031 (0.014) | 2.24 |
| | Friday | 0.154 (0.007) | 0.066 (0.013) | 5.15 |
| | Saturday | 0.124 (0.010) | 0.036 (0.014) | 2.51 |
| Friday | Tuesday | 0.070 (0.005) | -0.018 (0.014) | -1.29 |
| | Wednesday | 0.118 (0.008) | 0.031 (0.014) | 2.16 |
| | Thursday | 0.152 (0.011) | 0.065 (0.015) | 4.27 |
| | Friday | 0.154 (0.016) | 0.066 (0.019) | 3.54 |
| | Saturday | 0.124 (0.006) | - | - |
| Saturday | Tuesday | 0.075 (0.007) | -0.013 (0.015) | -0.87 |
| | Wednesday | 0.082 (0.006) | -0.005 (0.014) | -0.38 |
| | Thursday | 0.132 (0.008) | 0.044 (0.014) | 3.17 |
| | Friday | 0.151 (0.008) | 0.063 (0.013) | 4.69 |
| | Saturday | 0.173 (0.016) | 0.086 (0.017) | 4.89 |
| Sunday | Tuesday | 0.144 (0.023) | 0.056 (0.024) | 2.37 |
| | Wednesday | 0.081 (0.008) | -0.007 (0.015) | -0.44 |
| | Thursday | 0.101 (0.008) | 0.013 (0.015) | 0.89 |
| | Friday | 0.153 (0.013) | 0.065 (0.017) | 3.79 |
| | Saturday | 0.197 (0.012) | 0.109 (0.016) | 7.00 |

[a] Note that since DoW variables are binary, partial effects are always expressed as a change relative to an arbitrary reference case, which corresponds to the variable with missing PE.

Table 4.5. Partial effects by candidate DoW and DoW of purchase of Model LD-XDoW. OD-LD is taken at the value of 7 and the rest of variables are averaged. Parentheses show cluster robust standard errors.

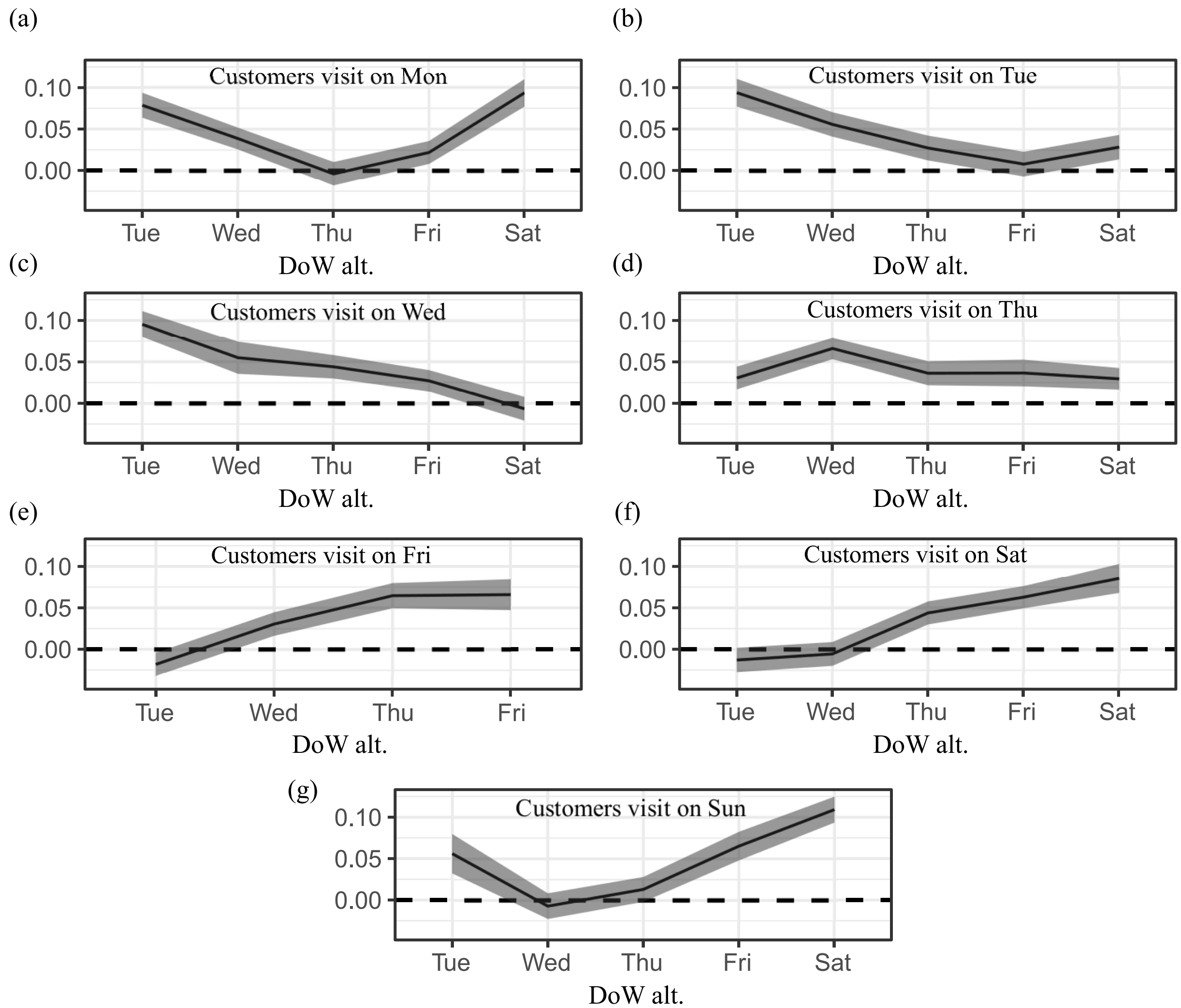


Figure 4.8. Partial effects by candidate DoW and DoW of purchase of Model LD-XDoW. OD-LD is taken at the value of 7 and the rest of variables are averaged. Error bars show cluster robust standard errors.

is notably higher than the empirical proportions observed. In contrast, in Model LD-XDoW, which incorporates the effect of OD-Lead-Time and crossed DoW, the predicted proportions become closer to the observed proportions. The two peaks are explained by the model and are the result of the combination of OD-Lead-Time and DoW effects. We further discuss this insight in Section 4.6.

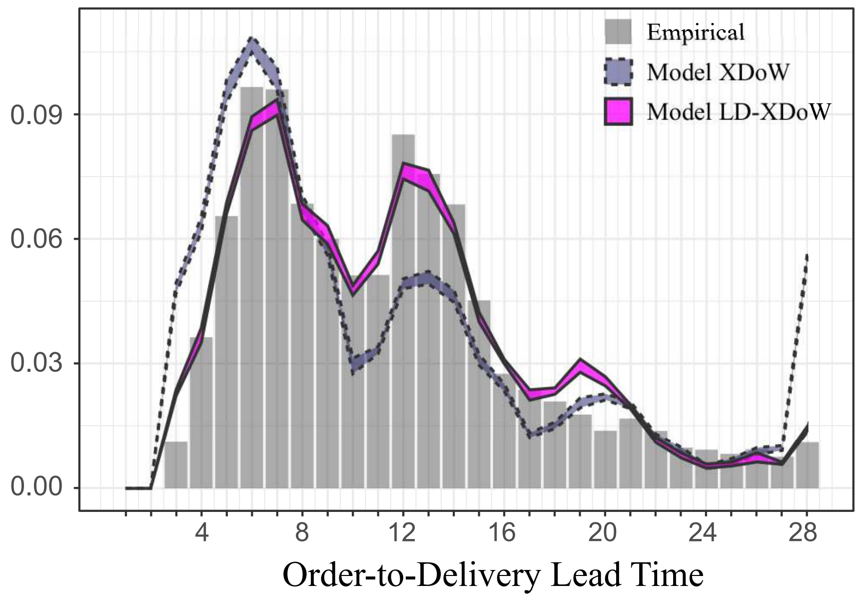


Figure 4.9. Predicted proportions. Results report the mean ± standard deviation of the proportions obtained of 5 simulations.

With a fairly simple behavioral model (the customer-retailer interaction model in Section 4.3.3) we have been able to understand the main features of the empirical proportions observed by the retail company. More importantly, we have been able to better understand the main drivers of customer choice when scheduling AHD. Our model and analysis provides several insights that can be used to more efficiently plan AHD operations. In the next section we discuss our empirical findings, and the managerial insights from our results.

## 4.6. Discussion and Operational Levers

In this section we discuss the results of Section 4.5. We divide our discussion in four subsections, each one focusing on a main insight obtained from our analysis.

### 4.6.1 Early Days are Not Usually Available for the Customers

In Figures 4.4 and 4.7 we show there is a significant and strong effect of OD-Lead-Time on customer availability. Availability is notably low for early candidate days that are close to the day of purchase, which supports Hypothesis 1. We attribute this effect to the fact that customers find the immediate future days unavailable due to family, work commitments or other. Another plausible explanation is that customers are more willing to reschedule commitments in later future periods to fit the AHD than to reschedule commitments in earlier periods. In other words, availability increases when the candidate day is later in the schedule because later days are not yet planned, or because plans are more easily rescheduled. Note that in Figures 4.4-(b) and 4.7-(b) the impact of time is stronger for early dates than for later dates. Thus, earlier candidate days are more sensitive to OD-Lead-Time than later days [3].

To test how customer choice changes with and without capacity constraints, in Figure 4.10, we show the simulated predicted proportions when all days are available for the retail company and are potentially offered, i.e., when there are no capacity constraints. When OD-Lead-Time of candidate days is taken into account in Model XDoW-LD, delivery dates earlier than one week are rarely chosen by customers even when they are offered.

To feasibly deliver the products at the customer's home, companies need to plan delivery capacity, such as delivery teams and vehicles, in advance [**2**]. Ample delivery capacity allows companies to reduce the minimum delivery lead time and to offer days with short OD-Lead-Time. However, this might come at high operational overage costs and

[3] A decreasing slope can be a attributed to using a logarithm function for OD-Lead-Time. When we used a linear function we found that likelihood significantly decreased, showing that a logarithm specification better captured customer behavior.

underutilized capacity when not planned accordingly with market demand and customer's expectations (see, for example, [**112**], [**26**], and [**40**]). We show that, in AHD, speed of delivery is of relative importance because availability plays a fundamental role in the final choice of the customers. This is particularly important for very early dates, in which customer availability is significantly low for most customers. Expensive operational efforts made by companies to achieve very-fast delivery are likely to have a tiny impact on customer experience because such dates are unlikely to be picked by customers when offered. Our results contrast with other studies that analyze short delivery lead-time operations and focus on the negative impact of increasing delivery lead times (e.g. [**3**], [**91**], [**133**], and [**38**]). In a different context, our results support other studies showing the relative importance of speed of access in non-urgent health care service systems (e.g. [**109**], [**108**], and [**83**]).

### 4.6.2 Day of Week of Purchase Reveals Day of Week Availability

When we incorporate the impact of candidate DoW and DoW of purchase we observe that, while there is a slight increase of likelihood (Table 4.2, 10th row, Model DoW and Model LD-DoW), in Figure 4.5-(c) and (d), the partial effects are roughly the same across all DoWs. We do not observe enough evidence to support that the candidate DoW nor the DoW of purchase has an impact on customer choice, (hypotheses 2 and 3.1).

In Figure 4.8, when we incorporate the crossed terms of DoW, we observe significant differences of candidate DoW availability depending on the DoW of purchase. For instance, availability when the candidate is Tuesday (Figure 4.8-(a)) is higher for customers that purchased the product on Mondays, Tuesdays or Wednesdays than those

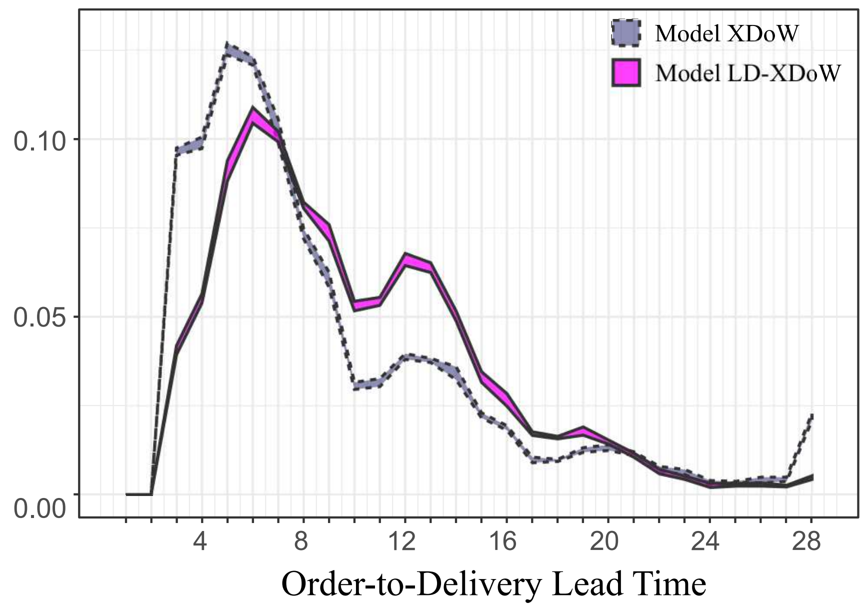


Figure 4.10. Predicted proportions without capacity constraints, i.e.; retailer filter contains all days in the schedule. Results report the mean ± standard deviation of the proportions obtained of 5 simulations.

that purchased the products on Fridays or Saturdays. Therefore, on top of the impact of OD-Lead-Time described in the previous section, DoW has a strong effect that depends on customer features, which we capture by incorporating the DoW crossed terms.

In Figure 4.9, the second peak in the observed and predicted proportions reveals that the impact of crossed DoW terms is strong even 10 days after purchase. We observe that a big portion of customers that made a purchase on a Saturday and were not able to schedule the delivery for the immediate weekend after purchase (first peak occurring around OD-Lead-Time 6-7) seem to be willing to wait one more week to schedule the delivery the following weekend (second peak occurring around OD-Lead-Time 12-13). Customers are selective about the DoW even if they have to wait more than one week.

Our results show that the DoW of purchase is a good indicator of the candidate DoW availability, supporting hypothesis 3.2. We attribute this result to the impact of weekly commitments such as work or school schedules. For instance, a customer who is not able to visit the retail outlet on a weekday due to work-related weekly availability is unlikely to schedule AHD for a weekday. This result supports previous studies focusing on the impact of DoW on customer choices in a variety of contexts such as travel mode, recreational activities and shopping trends (e.g. [**16**], [**141**], [**47**] , and [**144**]).

For the furniture retail company, the impact of DoW availability provides the supply chain team with information that can be used to more accurately forecast transportation needs and adjust delivery capacity in advance. For instance, our results show that Tuesday is likely to be available for customers purchasing on Tuesday, Wednesday and Thursday. In contrast, availability for Tuesdays of customers purchasing on Friday, Saturday or Sunday drops. Hence, by the end of Thursday, the retailer has an accurate estimate of the total amount of deliveries for the following Tuesday. A potential research direction can explore leveraging this information, as well as the sequential nature of DoW, by incorporating customer choice into capacity planning. The use of information to update demand forecasts is highly valuable and well studied in the literature (e.g. [**135**], [**54**], [**100**], and [**78**]).

### 4.6.3 Understanding the Observed Proportions

Our model allows us to understand the bi-modal shape of the empirical proportions. To this aim, we compare the results in Figure 4.9 and Figure 4.10, where we simulate the customer choice when all days have delivery capacity.

The occurrence of the two modes is a consequence of several effects taking place simultaneously:

The first peak occurs as a consequence of three factors: (1) Due to capacity constraints, very early dates (OD-Lead-Time 3-5) are often not offered (retailer availability filter). (2) Due to the impact of OD-Lead-Time, early dates are rarely chosen by customers due to availability even if they are offered (see Model XDoW-LD in Figure 4.10). (3) Due to DoW effects, the first day available for the customers is likely to be 5-6 days after purchase (see Model XDoW in Figure 4.10).

The second peak occurs as a consequence of a spill-over effect combined with the impact of DoW. Customers whose availability is strongly dominated by DoW can only visit the stores and schedule their delivery on certain days of week. For the sake of exposition, let customers who visit the store on weekends only schedule their delivery on weekends. For those customers, when the immediate next Saturday cannot be chosen, either due to capacity constraints or due to the impact of OD-Lead-Time, those customers need to wait until the following weekend, 13-14 days after purchase, around the occurrence of the second peak. The same can be argued, for example, for customers that visit the store and have availability only on Mondays, Tuesdays or Wednesdays.

Our numerical results in Figure 4.10 confirm our intuition about the second peak. When there are no capacity constraints and the OD-Lead-Time is not incorporated (Model XDoW), the second peak is barely noticeable. This is because most of the customers are able to schedule their delivery within the first week. When we incorporate OD-Lead-Time with no capacity constraints (Figure 4.10, Model XDoW-LD), the second peak significatively increases. In the presence of capacity constraints but no OD-Lead-Time effect (Figure 4.9, Model XDoW), the second peak is also noticeable. The second peak the highest with the combined effect OD-Lead-Time and capacity constraints (Figure 4.9 XDoW-LD).

To sum up, the bimodal shape of the empirical proportions observed by the furniture retail company is a consequence of a combined effect of capacity constraints, the impact of OD-Lead-Time, the crossed DoW and spill overs.

Understanding the drivers of customer choice that shape the empirical proportions can help the supply chain team of the retail company to plan efficiently AHD operations. While out of the scope of this chapter, our simple but insightful model can be used to easily simulate "what-ifs" scenarios to test delivery capacity allocation and scheduling policies. A research direction can consider embedding the customer-retailer interaction in a decision support system that plan delivery capacity by taking into account customer availability. Capacity planning with customer preferences has been shown to be valuable (see for example [**87**] and [**122**]), but few studies have focused on models that incorporate a customer availability.

### 4.6.4 Customer Availability Can be Predicted

One of the main concerns of the supply chain team of the furniture retail company is to use delivery capacity efficiently while improving customer satisfaction. Our framework shows how availability drives customer choice, and that the customer availability filter can be inferred from transactional data. This information can be used to increase customer satisfaction by strategically offering delivery dates.

For the sake of illustration, let us consider first an idealized scenario, in which the retail company knows in advance the availability of all customers. With that information, the retail company can offer delivery days such that all customers are maximally satisfied given the capacity constraints. This could be achieved, for example, by ensuring that customers wait the minimum possible time. For example, consider a customer that makes a purchase on a Saturday and only has availability on Saturdays. If the retailer has no delivery capacity for the first Saturday after purchase, the customer has no option but to wait two weeks for delivery. Had the retail company saved a spot on Saturday, the customer would have waited significantly less.

While the ideal scenario is impossible to implement because customer availability is not known in advance, our work shows that the retail company can infer customer availability. For instance, our results show that customers that make a purchase on Saturdays are likely to have lower availability on some weekdays. For the furniture retail company, saving a little delivery capacity on Fridays and Saturdays for customers that make a purchase on Saturdays can significantly increase customer satisfaction for those customers, while minimally affecting customer segments with higher availability. Therefore, not surprisingly, customers have strong DoW preferences, and our analysis shows that customers can be intuitively and easily segmented by DoW of purchase. This segmentation can be leveraged by the furniture company to customize delivery options with easy-to-implement policies and more efficiently use delivery capacity.

Several studies have shown that incorporating customer behavior in scheduling operations can be highly valuable (e.g., [**65**] and [**48**]). Our study constitutes a first step in showing how customer availability can be incorporated in AHD scheduling.

## 4.7. Conclusions

We focus on the last mile delivery scheduling operations of the furniture company that motivates our work in Chapters 2 and 3. We study customer choice of AHD date with a simple behavioral model that assumes that customers choose the earliest available date (available for the customer and available for the retailer). Our results suggest that in AHD, speed of delivery is of limited importance relative to customer availability. Customers of the furniture company rarely choose to schedule AHD on the dates immediately after the day of purchase, even if those dates are offered by the company. We show that customer availability increases with time, and it can be predicted based on the day of week the customers visited the store to make a purchase. Not surprisingly, our analysis reveals that customers have strong day of week preferences. Our study provides a method to quantify the effect of day of week on customer availability and its interaction with the effect of the order-to-delivery lead time (or speed of delivery). Our results suggest intuitive customer segmentation by day of the week of purchase, and easy-to-implement strategies to customize delivery capacity to these customer segments. The retail company can potentially improve its delivery operations by predicting and taking into account customer availability when planning delivery capacity.

CHAPTER 5

# Final Remarks

Motivated by the decision process of a major furniture manufacturer, this dissertation explores novel strategies to adapt middle mile and last mile operations of companies that commit to deliver their products to the customer's home. I show that the nature of home delivery and middle mile operations provides new opportunities for flexibility and adaptation that can reduce distribution costs across the retail supply chain.

The first part of this dissertation focuses on the impact of demand visibility, resulting from in-store orders and deliverable windows of made-to-order products, on policies for signing transportation procurement contracts with short temporal commitment. The second part of this dissertation incorporates network effects and shows, in contrast to existing studies of substitution effects, that companies can enhance the value of demand visibility if flexible transportation capacity is also available as an option. Finally, focusing on the last mile distribution, the third part of this dissertation studies customers' sensibility to delivery time windows in order to balance costly transportation capacity and customer satisfaction. The three parts reveal that in some cases order information at the store level can be used to more efficiently plan product distribution across the supply chain.

To conclude this dissertation, in the following sections I pose and discuss open questions related to the use of demand visibility with short term contracts, and questions on including customer behavior in last mile operations.

## 5.1. The Impact of Short Term Commitment Contracts Across the Supply Chain

New technological innovations are transforming the sourcing strategies across the supply chain. For example, in warehousing, Flexe, Stord, and Flow are companies that offer flexible on-demand warehousing capacity with short-term commitment contracts. In manufacturing, companies such as HP, Protolabs and Materialize use 3D printing to offer short-term flexible production capacity. In processing and manufacturing, there is a growing interest in mobile production capacity that is easily reconfigurable (e.g. [**14**] and [**88**]). [**45**] studies additional sources of flexibility across the supply chain.

As operations become more nimble, companies change their interaction with other actors in the supply chain. For example, in Chapter 2 and 3, we show that managers are more willing to commit to short-term contracts when they can use partial demand information. As a result, new questions emerge from a carrier perspective: Is there any benefit for a carrier in offering short-term commitment contracts? If so, when should carriers offer short commitment versus long commitment contracts and, at what cost? In dynamic environments where demand information is valuable, carriers can facilitate flexibility to adapt to such information through short-term contracts. Strategically managing and pricing information and flexibility can be valuable for both, carriers and shippers.

## 5.2. Leveraging Customer Availability and Preferences

The results in Chapter 4 identify managerial insights to plan more efficiently last mile distribution. The analysis of the transactional data from the furniture company reveals that speed of delivery is of relative importance because availability plays a fundamental role in the final choice of the customers. We also show that customer availability can be inferred from the data.

More work is needed to show how the company can leverage customer availability in attended home delivery. By predicting customer availability and strategically offering delivery dates, we suggest that the company can decrease last mile costs while decreasing customer waiting time. An interesting question arises: What is the minimum delivery capacity needed to ensure that customers wait at most one week on average? An interesting research direction can formally study this question with an assortment or a capacity planning model to find the minimum capacity needed to achieve a target service level (average customer waiting time) when customer choice is driven by availability. The study of scheduling policies with customer choice has gained attention in the past years (e.g. [**65**] and [**48**]). However, to the best of our knowledge, ours is the first model that considers customer availability as the main driver of customer choice.

Our results also show that the day of week of purchase is a good indicator of customer's availability. We suggest this information can be used as a form of advanced demand information when planning delivery capacity. For instance, by the end of Thursday, the company has a good estimate of the capacity needed for Tuesday of the following week. An interesting question emerges: How should the furniture company leverage such information? How can next week's estimates of last mile transportation needs be updated? And, how should delivery capacity be adjusted accordingly? Some studies have focused on integrating capacity planning with demand forecast updates (e.g. [**54**], [**101**] and [**134**]). Applying such models in attended home delivery scheduling can shed some light on these questions.

## 5.3. Sustainability of Home Delivery Services

In general terms, a big body of the literature has focused on increasing customer satisfaction by reducing delivery lead time as much as possible (e.g. [**70**], [**52**], [**104**], [**1**], [**38**] and [**39**]). However, questions about the sustainability of such supply chains remain unclear. For instance, some studies have found that home delivery is costly for companies and the environment (e.g., [**42**], [**104**], and [**37**]). Chapter 4 of this dissertation is one of the first studies highlighting that, in some situations, reducing delivery lead times below a threshold (offering, for example, options such as same-day delivery, or two-day delivery) translates to little benefits, if any. On the contrary, Chapters 2 and 3 show that the furniture company can benefit from using demand visibility resulting from longer delivery time windows of one or even two weeks.

As home delivery becomes a more and more popular option in retail, we should ask questions about how customers and retailers should manage delivery expectations. Should a company always offer indiscriminately same-day, next-day, or two-day delivery for all products to all customers? Do all customers have the same sensibility to delivery lead time? Can a company reduce transportation costs and carbon footprint impact by offering strategically options with shorter or longer delivery lead time? Answering these questions can provide insights to make retail supply chains more sustainable while meeting customer's expectations for home delivery of goods and services.

## My Last Remarks

In this dissertation, I borrowed tools and models from the transportation/logistics and the operations management literature. In Chapters 2 and 3, I applied results from inventory management to model distribution networks, and to unveil managerial insights hidden behind the complexity faced by the supply chain team of the furniture company. In Chapter 4, I applied ideas from the health care operations literature and school attainment models to study last mile operations of the furniture company. My research focuses on building intuition and managerial insights to identify novel strategies to deal with the complexity naturally present in the distribution networks. To do that I have merged two streams of the literature, transportation and operations management, which has allowed me to ask questions and tackle challenges from a different perspective. I am grateful that my advisors, my committee, and many other researchers from the transportation and the operations community have appreciated such a perspective. I am thankful for the help and support from these great communities.

## References

[1] Acimovic, J., Lim, M., and Mak, H.-Y. Moving beyond the speed-price trade-off. *MIT Sloan Management Review* (2018).

[2] Agatz, N., Campbell, A., Fleischmann, M., and Savelsbergh, M. Time slot management in attended home delivery. *Transportation Science 45*, 3 (2011), 435–449.

[3] Allon, G., Federgruen, A., and Pierson, M. How much is a reduction of your customers' wait worth? an empirical study of the fast-food drive-thru industry based on structural estimation methods. *Manufacturing & Service Operations Management 13*, 4 (2011), 489–507.

[4] Alp, O., Erkip, N. K., and Güllü, R. Outsourcing logistics: Designing transportation contracts between a manufacturer and a transporter. *Transportation Science 37*, 1 (2003), 23–39.

[5] Angelus, A., and Özer, Ö. Knowledge you can act on: optimal policies for assembly systems with expediting and advance demand information. *Operations Research 64*, 6 (2016), 1338–1371.

[6] Anupindi, R., and Bassok, Y. Supply contracts with quantity commitments and stochastic demand. In *Quantitative Models for Supply Chain Management*. Springer, 1999, pp. 197–232.

[7] Asdemir, K., Jacob, V. S., and Krishnan, R. Dynamic pricing of multiple home delivery options. *European Journal of Operational Research 196*, 1 (2009), 246–257.

[8] Atamtürk, A., and Hochbaum, D. S. Capacity acquisition, subcontracting, and lot sizing. *Management Science 47*, 8 (2001), 1081–1100.

[9] Atamtürk, A., and Küçükyavuz, S. Lot sizing with inventory bounds and fixed costs: Polyhedral study and computation. *Operations Research 53*, 4 (2005), 711–730.

[10] Banker, S. Last mile deliveries: Complex, costly, and critical. Forbes Magazine. Link: *https://www.forbes.com/sites/stevebanker/2020/08/26/last-mile-deliveries-complex-costly-and-critical/?sh=2f994c27723c*, Aug. 2020.

[11] Bassok, Y., and Anupindi, R. Analysis of supply contracts with total minimum commitment. *IIE Transactions 29*, 5 (1997), 373–381.

[12] Bassok, Y., and Anupindi, R. Analysis of supply contracts with commitments and flexibility. *Naval Research Logistics (NRL) 55*, 5 (2008), 459–477.

[13] BBC. Patients denied advanced bookings. CNN: *http://news.bbc.co.uk/1/hi/health/4112390.stm*, 2005.

[14] Becker, T., Lier, S., and Werners, B. Value of modular production concepts in future chemical industry production networks. *European Journal of Operational Research 276*, 3 (2019), 957–970.

[15] Bernstein, F., and DeCroix, G. A. Advance demand information in a multiproduct system. *Manufacturing & Service Operations Management 17*, 1 (2014), 52–65.

[16] Bertram, C., Meyerhoff, J., Rehdanz, K., and Wüstemann, H. Differences in the recreational value of urban parks between weekdays and weekends: A discrete choice analysis. *Landscape and Urban Planning 159* (2017), 5–14.

[17] Bhat, C. R., and Gossen, R. A mixed multinomial logit model analysis of weekend recreational episode type choice. *Transportation Research Part B: Methodological 38*, 9 (2004), 767–787.

[18] Boada-Collado, P., Chopra, S., Ibanez, M., and Smilowitz, K. Estimation of customer availability in attended home delivery. *Manuscript in preparation* (2020).

[19] Boada-Collado, P., Chopra, S., and Smilowitz, K. Partial demand information and commitment in dynamic transportation procurement. *Transportation Science 54*, 3 (2020), 588–605.

[20] Boada-Collado, P., Chopra, S., and Smilowitz, K. The value of information and flexibility with temporal commitments. *Second round revision in Manufacturing and Service Operations Management. Available at SSRN 3452915* (2020).

[21] Boada-Collado, P., and Martínez-de Albéniz, V. Estimating and optimizing the impact of inventory on consumer choices in a fashion retail setting. *Manufacturing & Service Operations Management 22*, 3 (2020), 582–597.

[22] Boyer, K. K., Prud'homme, A. M., and Chung, W. The last mile challenge: evaluating the effects of customer density and delivery window patterns. *Journal of Business Logistics 30*, 1 (2009), 185–201.

[23] Bruck, B. P., Cordeau, J.-F., and Iori, M. A practical time slot management and routing problem for attended home services. *Omega 81* (2018), 208–219.

[24] Buis, M. L. The consequences of unobserved heterogeneity in a sequential logit model. *Research in Social Stratification and Mobility 29*, 3 (2011), 247–262.

[25] Buschkühl, L., Sahling, F., Helber, S., and Tempelmeier, H. Dynamic capacitated lot-sizing problems: a classification and review of solution approaches. *Or Spectrum 32*, 2 (2010), 231–261.

[26] Cachon, G., and Terwiesch, C. *Matching supply with demand*. McGraw-Hill Publishing, 2008.

[27] Cameron, S. V., and Heckman, J. J. Life cycle schooling and dynamic selection bias: Models and evidence for five cohorts of american males. *Journal of Political economy 106*, 2 (1998), 262–333.

[28] Campbell, A. M., and Savelsbergh, M. W. Decision support for consumer direct grocery initiatives. *Transportation Science 39*, 3 (2005), 313–327.

[29] Caro, F., and Martínez-de Albéniz, V. Fast fashion: Business model overview and research opportunities. In *Retail supply chain management*. Springer, 2015, pp. 237–264.

[30] Chen, F. Market segmentation, advanced demand information, and supply chain performance. *Manufacturing & Service Operations Management 3*, 1 (2001), 53–67.

[31] Chod, J., Markakis, M., and Trichakis, N. On the learning benefits of resource flexibility.

[32] Chod, J., and Rudi, N. Resource flexibility with responsive pricing. *Operations Research 53*, 3 (2005), 532–548.

[33] Chod, J., Rudi, N., and Van Mieghem, J. A. Operational flexibility and financial hedging: Complements or substitutes? *Management Science 56*, 6 (2010), 1030–1045.

[34] Chod, J., and Zhou, J. Resource flexibility and capital structure. *Management Science 60*, 3 (2013), 708–729.

[35] Chuang, H. H.-C., Oliva, R., and Perdikaki, O. Traffic-based labor planning in retail stores. *Production and Operations Management 25*, 1 (2016), 96–113.

[36] Copil, K., Wörbelauer, M., Meyr, H., and Tempelmeier, H. Simultaneous lotsizing and scheduling problems: a classification and review of models. *OR spectrum 39*, 1 (2017), 1–64.

[37] Crandall, R. Eliminating last-mile environmental inefficiencies. Link: *https://www.ascm.org/ascm-insights/eliminating-last-mile-environmental-inefficiencies/*, Dec. 2020.

[38] Cui, R., Li, M., and Li, Q. Value of high-quality logistics: Evidence from a clash between sf express and alibaba. *Management Science 66*, 9 (2020), 3879–3902.

[39] Cui, R., Sun, T., Lu, Z., and Golden, J. Sooner or later? promising delivery speed in online retail. *Promising Delivery Speed in Online Retail (March 29, 2020)* (2020).

[40] De Treville, S., Bicer, I., Chavez-Demoulin, V., Hagspiel, V., Schürhoff, N., Tasserit, C., and Wager, S. Valuing lead time. *Journal of Operations Management 32*, 6 (2014), 337–346.

[41] De Vries, J., Roy, D., and De Koster, R. Worth the wait? how restaurant waiting time influences customer behavior and revenue. *Journal of operations Management 63* (2018), 59–78.

[42] Deloison, T., Hannon, E., Huber, A., Heid, B., Klink, C., Sahay, R., and Wolff, C. World economic forum report: The future of the last-mile ecosystem. *http://www3.weforum.org/docs/WEF_Future_of_the_last_mile_ecosystem.pdf*, 2020.

[43] Dixon, S., Sampson, F. C., O'Cathain, A., and Pickin, M. Advanced access: more than just gp waiting times? *Family practice 23*, 2 (2006), 233–239.

[44] Dolan, S. The challenges of last mile delivery logistics and the tech solutions cutting costs in the final mile. Business Insider. Link: *https://www.businessinsider.com/last-mile-delivery-shipping-explained*, Jan 2021.

[45] Dolgui, A., Ivanov, D., and Sokolov, B. Reconfigurable supply chain: the x-network. *International Journal of Production Research* (2020), 1–26.

[46] Dong, L., Kouvelis, P., and Su, P. Global facility network design in the presence of competition. *European Journal of Operational Research 228*, 2 (2013), 437–446.

[47] Ehrenthal, J., Honhon, D., and Van Woensel, T. Demand seasonality in retail inventory management. *European Journal of Operational Research 238*, 2 (2014), 527–539.

[48] Feldman, J., Liu, N., Topaloglu, H., and Ziya, S. Appointment scheduling under patient preference and no-show behavior. *Operations Research 62*, 4 (2014), 794–811.

[49] Ferguson, M. E. When to commit in a serial supply chain with forecast updating. *Naval Research Logistics (NRL) 50*, 8 (2003), 917–936.

[50] Ferguson, M. E., DeCroix, G. A., and Zipkin, P. H. Commitment decisions with partial information updating. *Naval Research Logistics (NRL) 52*, 8 (2005), 780–795.

[51] Fine, C. H., and Freund, R. M. Optimal investment in product-flexible manufacturing capacity. *Management Science 36*, 4 (1990), 449–466.

[52] Fisher, M. L., Gallino, S., and Xu, J. J. The value of rapid delivery in omnichannel retailing. *Journal of Marketing Research 56*, 5 (2019), 732–748.

[53] Freedman, D. A. On the so-called huber sandwich estimator and robust standard errors. *The American Statistician 60*, 4 (2006), 299–302.

[54] Gallego, G., and Özer, Ö. Integrating replenishment decisions with advance demand information. *Management science 47*, 10 (2001), 1344–1360.

[55] Gallino, S., Moreno, A., and Stamatopoulos, I. Channel integration, sales dispersion, and inventory management. *Management Science 63*, 9 (2017), 2813–2831.

[56] Garrido, R. A. Procurement of transportation services in spot markets under a double-auction scheme with elastic demand. *Transportation Research Part B: Methodological 41*, 9 (2007), 1067–1078.

[57] Glaeser, C. K., Fisher, M., and Su, X. Optimal retail location: Empirical methodology and application to practice. *Manufacturing & Service Operations Management 21*, 1 (2019), 86–102.

[58] Goldman, D. Amazon shares soar on fourth straight profitable quarter. CNN: *https://money.cnn.com/2016/04/28/technology/amazon-earnings//*, Apr. 2016.

[59] Goyal, M., and Netessine, S. Strategic technology choice and capacity investment under demand uncertainty. *Management Science 53*, 2 (2007), 192–207.

[60] Goyal, M., and Netessine, S. Volume flexibility, product flexibility, or both: The role of demand correlation and product substitution. *Manufacturing & service operations management 13*, 2 (2011), 180–193.

[61] Granados, N., Gupta, A., and Kauffman, R. J. Online and offline demand and price elasticities: Evidence from the air travel industry. *Information Systems Research 23*, 1 (2012), 164–181.

[62] Graves, S. C., and Tomlin, B. T. Process flexibility in supply chains. *Management Science 49*, 7 (2003), 907–919.

[63] Greene, W. H., and Hensher, D. A. *Modeling ordered choices: A primer.* Cambridge University Press, 2010.

[64] Gupta, D., and Denton, B. Appointment scheduling in health care: Challenges and opportunities. *IIE transactions 40*, 9 (2008), 800–819.

[65] Gupta, D., and Wang, L. Revenue management for a primary-care clinic in the presence of patient choice. *Operations Research 56*, 3 (2008), 576–592.

[66] Haksöz, Ç., and Seshadri, S. Supply chain operations in the presence of a spot market: a review with discussion. *Journal of the Operational Research Society 58*, 11 (2007), 1412–1429.

[67] Harrison, J. M., and Van Mieghem, J. A. Multi-resource investment strategies: Operational hedging under demand uncertainty. *European Journal of Operational Research 113*, 1 (1999), 17–29.

[68] Hogg, R. V., McKean, J., and Craig, A. T. *Introduction to mathematical statistics*. Pearson Education, 2005.

[69] Hopp, W. J., Tekin, E., and Van Oyen, M. P. Benefits of skill chaining in serial production lines with cross-trained workers. *Management Science 50*, 1 (2004), 83–98.

[70] Howard, R. Same-day delivery: A checklist for retailers seeking an antidote to amazon: Part one. *https://parcelindustry.com/article-4003-Same-Day-Delivery-A-Checklist-for-Retailers-Seeking-an-Antidote-to-Amazon-Part-One.html*, 2014.

[71] Ibanez, M. R., Clark, J. R., Huckman, R. S., and Staats, B. R. Discretionary task ordering: Queue management in radiological services. *Management Science 64*, 9 (2018), 4389–4407.

[72] Ibanez, M. R., and Toffel, M. W. How scheduling can bias quality assessment: Evidence from food-safety inspections. *Management Science 66*, 6 (2020), 2396–2416.

[73] Inderfurth, K., Kelle, P., and Kleber, R. Dual sourcing using capacity reservation and spot market: Optimal procurement policy and heuristic parameter determination. *European Journal of Operational Research 225*, 2 (2013), 298–309.

[74] Ito, K. Do consumers respond to marginal or average price? evidence from non-linear electricity pricing. *American Economic Review 104*, 2 (2014), 537–63.

[75] Jain, A. Responding to shipment delays: The roles of operational flexibility & lead-time visibility. *Decision Sciences 49*, 2 (2018), 306–334.

[76] Jordan, W. C., and Graves, S. C. Principles on the benefits of manufacturing process flexibility. *Management Science 41*, 4 (1995), 577–594.

[77] Karaesmen, F. Value of advance demand information in production and inventory systems with shared resources. In *Handbook of Stochastic Models and Analysis of Manufacturing System Operations*. Springer, 2013, pp. 139–165.

[78] Karaesmen, F., Liberopoulos, G., and Dallery, Y. The value of advance demand information in production/inventory systems. *Annals of Operations Research 126*, 1-4 (2004), 135–157.

[79] Kay, M. G., and Warsing, D. P. Estimating ltl rates using publicly available empirical data. *International Journal of Logistics: Research and Applications 12*, 3 (2009), 165–193.

[80] King, G., and Roberts, M. E. How robust standard errors expose methodological problems they do not fix, and what to do about it. *Political Analysis* (2015), 159–179.

[81] Klein, R., Neugebauer, M., Ratkovitch, D., and Steinhardt, C. Differentiated time slot pricing under routing considerations in attended home delivery. *Transportation Science 53*, 1 (2017), 236–255.

[82] Lim, S., Gao, F., and Tan, T. Slow and sure, or fast and furious? an empirical study about omnichannel demand sensitivity to fulfillment lead time. *An Empirical Study About Omnichannel Demand Sensitivity to Fulfillment Lead Time (February 21, 2020). SMU Cox School of Business Research Paper Forthcoming* (2020).

[83] Liu, N., Finkelstein, S. R., Kruk, M. E., and Rosenthal, D. When waiting to see a doctor is less irritating: Understanding patient preferences and choice behavior in appointment scheduling. *Management Science 64*, 5 (2018), 1975–1996.

[84] Liu, R., Xie, X., and Garaix, T. Hybridization of tabu search with feasible and infeasible local searches for periodic home health care logistics. *Omega 47* (2014), 17–32.

[85] Lu, L. X., and Van Mieghem, J. A. Multimarket facility network design with offshoring applications. *Manufacturing & Service Operations Management 11*, 1 (2009), 90–108.

[86] Mackert, J. Choice-based dynamic time slot management in attended home delivery. *Computers & Industrial Engineering 129* (2019), 333–345.

[87] Mahajan, S., and Van Ryzin, G. Stocking retail assortments under dynamic consumer substitution. *Operations Research 49*, 3 (2001), 334–351.

[88] Malladi, S. S., Erera, A. L., and White III, C. C. A dynamic mobile production capacity and inventory control problem. *IISE Transactions* (2020), 1–18.

[89] Mare, R. D. Social background and school continuation decisions. *Journal of the american statistical association 75*, 370 (1980), 295–305.

[90] Mare, R. D. Change and stability in educational stratification. *American sociological review* (1981), 72–87.

[91] Marino, G., Zotteri, G., and Montagna, F. Consumer sensitivity to delivery lead time: a furniture retail case. *International Journal of Physical Distribution & Logistics Management* (2018).

[92] Martínez-Costa, C., Mas-Machuca, M., Benedito, E., and Corominas, A. A review of mathematical programming models for strategic capacity planning in manufacturing. *International Journal of Production Economics 153* (2014), 66–85.

[93] Martínez-de Albéniz, V., and Simchi-Levi, D. A portfolio approach to procurement contracts. *Production and Operations Management 14*, 1 (2005), 90–114.

[94] Mazareanu, E. Number of parcels distributed worldwide from 2015 to 2019, by sector. *https://www.statista.com/statistics/737418/parcel-traffic-worldwide-by-sector/*, Sept. 2020.

[95] Mood, C. Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European sociological review 26*, 1 (2010), 67–82.

[96] Mood, C. Logistic regression: Uncovering unobserved heterogeneity. Tech. rep., Working Paper. http://www. su. se/polopoly_fs/1.341160. 1501927239, 2017.

[97] Naim, M. M., Potter, A. T., Mason, R. J., and Bateman, N. The role of transport flexibility in logistics provision. *The International Journal of Logistics Management 17*, 3 (2006), 297–311.

[98] NRP. 2016 retail holiday planning playbook. National Retail Federation, June 2016.

[99] Osadchiy, N., and Kc, D. Are patients patient? the role of time to appointment in patient flow. *Production and Operations Management 26*, 3 (2017), 469–490.

[100] Özer, Ö. Replenishment strategies for distribution systems under advance demand information. *Management Science 49*, 3 (2003), 255–272.

[101] Özer, Ö., and Wei, W. Inventory control with limited capacity and advance demand information. *Operations Research 52*, 6 (2004), 988–1000.

[102] Özkaya, E., Keskinocak, P., Joseph, V. R., and Weight, R. Estimating and benchmarking less-than-truckload market rates. *Transportation Research Part E: Logistics and Transportation Review 46*, 5 (2010), 667–682.

[103] Papier, F. Supply allocation under sequential advance demand information. *Operations Research 64*, 2 (2016), 341–361.

[104] Pearce, F. In store or online - what's the enviromentally friendliest way to shop? *https://ensia.com/features/environmental-cost-online-shopping-delivery*, 2019.

[105] Pickin, M., O'Cathain, A., Sampson, F. C., and Dixon, S. Evaluation of advanced access in the national primary care collaborative. *British Journal of General Practice 54*, 502 (2004), 334–340.

[106] Prasad, A., Stecke, K. E., and Zhao, X. Advance selling by a newsvendor retailer. *Production and Operations Management 20*, 1 (2011), 129–142.

[107] Raftery, A. E., and Hout, M. Maximally maintained inequality: Expansion, reform, and opportunity in irish education, 1921-75. *Sociology of education* (1993), 41–62.

[108] Rubin, G., Bate, A., George, A., Shackley, P., and Hall, N. Preferences for access to the gp: a discrete choice experiment. *British Journal of General Practice 56*, 531 (2006), 743–748.

[109] Sampson, F., Pickin, M., O'Cathain, A., Goodall, S., and Salisbury, C. Impact of same-day appointments on patient satisfaction with general practice appointment systems. *British Journal of General Practice 58*, 554 (2008), 641–643.

[110] Sargut, F. Z., and Romeijn, H. E. Capacitated production and subcontracting in a serial supply chain. *IIE Transactions 39*, 11 (2007), 1031–1043.

[111] Shen, B., Choi, T.-M., and Minner, S. A review on supply chain contracting with information considerations: information updating and information asymmetry. *International Journal of Production Research* (2018), 1–39.

[112] So, K. C. Price and time competition for service delivery. *Manufacturing & Service Operations Management 2*, 4 (2000), 392–409.

[113] Song, J., and Regan, A. Approximation algorithms for the bid construction problem in combinatorial auctions for the procurement of freight transportation contracts. *Transportation Research Part B: Methodological 39*, 10 (2005), 914–933.

[114] Srinivasan, R., and Swink, M. An investigation of visibility and flexibility as complements to supply chain analytics: An organizational information processing theory perspective. *Production and Operations Management 27*, 10 (2018), 1849–1867.

[115] SteadieSeifi, M., Dellaert, N. P., Nuijten, W., Van Woensel, T., and Raoufi, R. Multimodal freight transportation planning: A literature review. *European Journal of Operational Research 233*, 1 (2014), 1–15.

[116] Stock, R. M., von Hippel, E., and Gillert, N. L. Impacts of personality traits on consumer innovation success. *Research Policy 45*, 4 (2016), 757–769.

[117] Sull, D., and Turconi, S. Fast fashion lessons. *Business Strategy Review 19*, 2 (2008), 4–11.

[118] Tan, T. F., and Netessine, S. When does the devil make work? an empirical study of the impact of workload on worker productivity. *Management Science 60*, 6 (2014), 1574–1593.

[119] Tibben-Lembke, R. S. N-period contracts with ordering constraints and total minimum commitments: optimal and heuristic solutions. *European Journal of Operational Research 156*, 2 (2004), 353–374.

[120] Tjokroamidjojo, D., Kutanoglu, E., and Taylor, G. D. Quantifying the value of advance load information in truckload trucking. *Transportation Research Part E: Logistics and Transportation Review 42*, 4 (2006), 340–357.

[121] Tomlin, B., and Wang, Y. On the value of mix flexibility and dual sourcing in unreliable newsvendor networks. *Manufacturing & Service Operations Management 7*, 1 (2005), 37–57.

[122] Topaloglu, H. Joint stocking and product offer decisions under the multinomial logit model. *Production and Operations Management 22*, 5 (2013), 1182–1199.

[123] Topan, E., Tan, T., Van Houtum, G.-J., and Dekker, R. Using imperfect advance demand information in lost-sales inventory systems with the option of returning inventory. *IISE Transactions 50*, 3 (2018), 246–264.

[124] Train, K. E. *Discrete choice methods with simulation.* Cambridge university press, 2009.

[125] Truong, V.-A. Optimal advance scheduling. *Management Science 61*, 7 (2015), 1584–1597.

[126] Tsai, M.-T., Saphores, J.-D., and Regan, A. Valuation of freight transportation contracts under uncertainty. *Transportation Research Part E: Logistics and Transportation Review 47*, 6 (2011), 920–932.

[127] Urban, T. L. Supply contracts with periodic, stationary commitment. *Production and Operations Management 9*, 4 (2000), 400–413.

[128] Van Mieghem, J. A. Investment strategies for flexible resources. *Management Science 44*, 8 (1998), 1071–1078.

[129] Van Mieghem, J. A. Risk mitigation in newsvendor networks: Resource diversification, flexibility, sharing, and hedging. *Management Science 53*, 8 (2007), 1269–1288.

[130] Van Mieghem, J. A. V., and Rudi, N. Newsvendor networks: Inventory management and capacity investment with discretionary activities. *Manufacturing & Service Operations Management 4*, 4 (2002), 313–335.

[131] van Ryzin, G., and Vulcano, G. A market discovery algorithm to estimate a general class of nonparametric choice models. *Management Science 61*, 2 (2015), 281–300.

[132] Viet, N. Q., Behdani, B., and Bloemhof, J. The value of information in supply chain decisions: A review of the literature and research agenda. *Computers & Industrial Engineering 120* (2018), 68–82.

[133] Voccia, S. A., Campbell, A. M., and Thomas, B. W. The same-day delivery problem for online purchases. *Transportation Science 53*, 1 (2019), 167–184.

[134] Wang, T., Atasu, A., and Kurtuluş, M. A multiordering newsvendor model with dynamic forecast evolution. *Manufacturing & Service Operations Management 14*, 3 (2012), 472–484.

[135] Wang, T., and Toktay, B. L. Inventory management with advance demand information and flexible delivery. *Management Science 54*, 4 (2008), 716–732.

[136] Wang, W.-Y., and Gupta, D. Adaptive appointment systems with patient preferences. *Manufacturing & Service Operations Management 13*, 3 (2011), 373–389.

[137] Weiler, W. C. A sequential logit model of the access effects of higher education institutions. *Economics of Education Review 5*, 1 (1986), 49–55.

[138] Wieberneit, N. Service network design for freight transportation: a review. *OR spectrum 30*, 1 (2008), 77–112.

[139] Wooldridge, J. M. *Econometric analysis of cross section and panel data*. MIT press, 2010.

[140] Wooldridge, J. M. *Introductory econometrics: A modern approach.* Nelson Education, 2016.

[141] Yang, L., Shen, Q., and Li, Z. Comparing travel mode and trip chain choices between holidays and weekdays. *Transportation Research Part A: Policy and Practice 91* (2016), 273–285.

[142] Yang, X., and Strauss, A. K. An approximate dynamic programming approach to attended home delivery management. *European Journal of Operational Research 263*, 3 (2017), 935–945.

[143] Yang, X., Strauss, A. K., Currie, C. S., and Eglese, R. Choice-based demand management and vehicle routing in e-fulfillment. *Transportation science 50*, 2 (2014), 473–488.

[144] Yun, D.-S., and O'Kelly, M. E. Modeling the day-of-the-week shopping activity and travel patterns. *Socio-Economic Planning Sciences 31*, 4 (1997), 307–319.

[145] Zolfagharinia, H., and Haughton, M. The benefit of advance load information for truckload carriers. *Transportation Research Part E: Logistics and Transportation Review 70* (2014), 34–54.

[146] Zolfagharinia, H., and Haughton, M. Effective truckload dispatch decision methods with incomplete advance load information. *European Journal of Operational Research 252*, 1 (2016), 103–121.

[147] Zuidwijk, R. A., and Veenstra, A. W. The value of information in container transport. *Transportation Science 49*, 3 (2014), 675–685.

APPENDIX A

# Appendix from Chapter 2

## A.1. Proofs

Proof of Lemma 1. Both arguments of the $min\{\bullet,\bullet\}$ function of $V$ are increasing with $d$. Assume that for some $d'$ the contract option is optimal. Then: $C^0(d') + \delta\,\mathbb{V} \geq C^1(d') + \Delta(\tau)\,\mathbb{E}_{D^w}[C^1(k(d), D^w)] + \delta^\tau\,\mathbb{V}$. Now consider $d'' = d' + \theta$ for $\theta \geq 0$. Then, $C^0(d'') + \delta\,\mathbb{V} = c'd' + c'\theta + \delta\,\mathbb{V} = C^0(d') + \delta\,\mathbb{V} + c'\theta \geq C^1(d') + \Delta(\tau)\,\mathbb{E}_{D^w}[C^1(k(d), D^w)] + \delta^\tau\,\mathbb{V} + c'\theta \geq C^1(d'') + \Delta(\tau)\,\mathbb{E}_{D^w}[C^1(k(d), D^w)] + \delta^\tau\,\mathbb{V}$.

The last inequality is true because the contract is Efficient. Similar argument can be followed to show that if for $d'$ the outsourcing option is optimal, then it is also optimal for $d'' = d' - \theta$.

□

Proof of Proposition 1. We prove by contradiction: Let the inequality be true and assume that, for some demand, it is optimal to sign the contract. By monotonicity proven in lemma (1) and the fact that $C^1(0) > 0$ and $C^0(0) = 0$, there exists an optimal threshold policy with threshold $\hat{d}$. When the demand assumes the value of the threshold, the two options are equally cost-efficient:

$$C^0(\hat{d}) + \delta\,\mathbb{V} = C^1(\bar{k}, \hat{d}) + \Delta(\tau)\,\mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] + \delta^\tau\,\mathbb{V}\,. \tag{A.1}$$

We take the expectations of equation (2.4) with respect to $D^w$:

(A.2)

$$\mathbb{V} = \mathbb{E}_{D^w}[V(D^w)] = \int_0^{\hat{d}}(C^0(d) + \delta\,\mathbb{V})dF + \int_{\hat{d}}^{\infty}(C^1(\bar{k}, d) + \Delta(\tau)\,\mathbb{E}_{D^w}[C^1(c, D^w)] + \delta^\tau\,\mathbb{V})dF =$$

$$= \mathbb{C}^0(\hat{d}) + \delta\,\mathbb{V}\,F(\hat{d}) + \mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] - \mathbb{C}^1(\bar{k}, \hat{d}) + (\Delta(\tau)\,\mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] + \delta^\tau\,\mathbb{V})(1 - F(\hat{d})) =$$

$$= \mathbb{C}^0(\hat{d}) + \delta\,\mathbb{V}\,F(\hat{d}) + \mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] - \mathbb{C}^1(\bar{k}, \hat{d}) + (C^0(\hat{d}) - C^1(\bar{k}, \hat{d}) + \delta\,\mathbb{V})(1 - F(\hat{d})),$$

where the last equality is obtained by using equation (A.1). Therefore, if there exist a threshold, then it is such that it must solve the system of equations (A.2) and (A.1).

We now show that the system has only one solution for $\hat{d}$. First, divide the first equation of (2) by $\Delta(\tau)$ and substitute $\mathbb{V}$ in the second equation. By rearranging terms, we obtain the equation shown in Proposition 2:

$$TF(\hat{d}) = 0,$$

where $TF(x)$ is the 'Threshold Function':

$$TF(x) = \mathbb{C}^0(x) - \mathbb{C}^1(\bar{k}, x) + (1 - F(x) + \tfrac{1}{\Delta(\tau)})(C^0(x) - C^1(\bar{k}, x)).$$

We split the analysis of $TF$ into two cases:

If $\hat{d} < \bar{k}$, by taking the first and second derivative we can see that $TF$ is increasing and concave:

$$\tfrac{d(TF(x))}{dx}|_{x<\bar{k}} = (1 - F(x) + \tfrac{1}{\Delta(\tau})c' > 0,$$
$$\tfrac{d^2(TF(x))}{dx^2}|_{x<\bar{k}} = -f(x)c' \leq 0.$$

If $\hat{d} \geq \bar{k}$, the $TF(x)$ is constant:

$$TF(x)|_{x \geq \bar{k}} = c' \int_0^{\bar{k}} ddF + kc'(F(x) - F(\bar{k})) - c\bar{k}F(x) + (1 - F(x) + \tfrac{1}{\Delta(\tau)})(\bar{k}(c - c')) =$$
$$= c' \int_0^{\bar{k}} ddF - \bar{k}c'F(\bar{k}) + \bar{k}(c - c') + \tfrac{\bar{k}(c-c')}{\Delta(\tau)} = \mathbb{E}_{D^w}[C^0(D^w)] - \mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] + \tfrac{\bar{k}(c'-c)}{\Delta(\tau)}.$$

where the last equality is obtained by adding and subtracting $\int_{\bar{k}}^{\infty} ddF$ and rearranging terms.

Note that the threshold is obtained at the point $x$ where $TF(x) = 0$. Because $TF$ is constant when $x \geq \bar{k}$, if such a point is not reached in the range of $x < \bar{k}$, then it is not reached in all the domain of $x$. Thus, there exists a solution if and only if $\mathbb{E}_{D^w}[C^0(D^w)] - \mathbb{E}_{D^w}[C^1(\bar{k}, D^w)] + \frac{\bar{k}(c'-c)}{\Delta(\tau)} < 0$ (i.e. when the CCBC is held), as stated in the proposition. Hence, if the condition of the lemma holds, the system (2) is not feasible and, therefore, an optimal threshold policy does not exist. Therefore, it contradicts our assumption.

Finally, as the condition implies that the expected cost of the contract mode is larger than outsourcing and by Lemma 1 the policy has to be monotonic, it is always optimal to outsource the whole demand. □

Proof of Proposition 2. The result is a direct consequence of Lemma 1 and Proposition 1. □

Proof of Proposition 3. The cost of the contract option $C^1(k_v, d)$ today and after $\tau$ periods is convex and can be formulated as follows:

$$C^{1,\tau}(k_v,d)=\begin{cases} f_l(k_v,d):=ck_v+c'(d-k_v)+\Delta(\tau)(ck_v-c'k_v(1-F(k_v))+c'\int\limits_{k_v}^{\infty}ddF) & \text{if } d<k_v \\ f_u(k_v,d):=ck_v+\Delta(\tau)(ck_v-c'k_v(1-F(k_v))+c'\int\limits_{k_v}^{\infty}ddF) & \text{if } d\geq k_v. \end{cases}$$

Note that $f_l(k_v,d)$ and $f_u(k_v,d)$ have their unique minimizers at $k_l^*=F^{-1}((\frac{c'-c}{c'}-\frac{c}{c'\Delta(\tau)})$ and $k_u^*=F^{-1}(\frac{c'-c}{c'}-\frac{c}{c'\Delta(\tau)}+\frac{1}{\Delta(\tau)})$ respectively and $k_l^*<k_u^*$. If $d<k_l^*$ ($d>k_u^*$), $f_l(k_v,d)$ ($f_l(k_v,d)$) has its minimizer active $k_l^*$. On the other hand, when $k_l^*\leq d<k_u^*$, the minimizer is the intersection of the two functions, $k^*=d$. □

Proof of Lemma 2.

$$\mathbb{E}_{D^w}[C^1(k_u^*,D^w)]-\mathbb{E}_{D^w}[C^0(D^w)]=k_u^*c+c'\int\limits_{k_u^*}^{\infty}ddF-c'k_u^*(1-F(k_u^*))-c'\int\limits_{0}^{\infty}ddF=$$
$$=k_u^*(c-c')+k_u^*c'F(k_u^*)-c'\int\limits_{0}^{k_u^*}ddF=\frac{k_u^*(c'-c)}{\Delta(\tau)}-c'\int\limits_{0}^{k_u^*}ddF<\frac{k_u^*(c'-c)}{\Delta(\tau)},$$

where the last equality is obtained by using $F(k_u^*)=\frac{c'-c}{c'}-\frac{c}{c'\Delta(\tau)}+\frac{1}{\Delta(\tau)}$ and canceling terms.

□

Proof of Proposition 4. The proof follows the same steps of Proposition 1 but with a capacity that depends on the demand seen during the first period. First we show that the contract mode is Efficient:

**Lemma 3.** *The contract mode when a single capacity contract is chosen from among a continuum set of capacities is Efficient.*

**Proof.** We show the result for the three different cases. If $x \geq k_l^*$, $k^*(x) = k_l^*$, which is constant:

$$\tfrac{d(C^{1,\tau}(k(x),x)}{dx} = \tfrac{d}{dx}(ck_l^* + \Delta(\tau))\,\mathbb{E}_{D^w}[C^1(k_l^*, D^w)]) = 0.$$

If $x \geq k_u^*$, $k^*(x) = k_u^*$, which is constant:

$$\tfrac{d(C^{1,\tau}(k(x),x)}{dx} = \tfrac{d}{dx}(ck_u^* + c'[x - k_u^*]^+ + \Delta(\tau))\,\mathbb{E}_{D^w}[C^1(k_u^*, D^w)]) = c'.$$

If $k_l^* \leq x \leq k_u^*$, $k^*(x) = x$:

(A.3)

$$\tfrac{d(C^{1,\tau}(k(x),x)}{dx} = \tfrac{d}{dx}(cx + \Delta(\tau))\,\mathbb{E}_{D^w}[C^1(x, D^w)]) = c + \Delta(\tau)(c - c'(1 - F(x))) \leq c'.$$

The last inequality in equation (A.3) follows from the fact that the derivative is always increasing and the maximum is reached at the point $x = k_u^*$ with a value of $c'$. $\square$

Hence, from lemma (3) the optimal policy is monotonic. In contrast with Proposition 1, the contract option is $C(k(d), d)$, which is a piece-wise linear function, and $\beta(d)$ is a function that depends explicitly on the next period demand, which complicates slightly the algebra. The proof uses the results in Lemma 1, Lemma 2, Lemma 3 and Proposition 3 .

We skip the rest of the proof. $\square$

Proof of Theorem 1. The results found in sections (2.3.2) and (2.3.3) (and the section in the appendix A.3 about Commitment under discrete set of capacities) show that for any set of real values that form contract capacities (one element, discrete or a continuous set), there exist a compact subset of Useful Capacities that will be used if the decision is optimal. We have shown that in all possible cases the choice of the optimal capacity implied that the contract mode was Efficient if $c' > c$. Therefore, by lemma 1 the policy is monotonic. $\square$

Proof of Proposition 5. The cost of the contract option $C_t^1(1, d_t)$ today and after $\tau$ periods at time $t$ is convex and can be formulated as follows:

$$C_t^{1,\tau}(k_v, d_t) = \begin{cases} f_{l,t}(k_v, d_t) := -(c' - c)k_v + c'd_t + \sum_{i=1}^{\tau-1} \delta^i \mathbb{C}_{t+i}^1(k_v) & \text{if } d_t < k_v \\ f_{u,t}(k_v, d_t) := ck_v + \sum_{i=1}^{\tau-1} \delta^i \mathbb{C}_{t+i}^1(k_v) & \text{if } d_t \geq k_v. \end{cases}$$

Note that $f_{l,t}(k_v, d_t)$ and $f_{u,t}(k_v, d_t)$ have their unique minimizers at $k_{l,t}^*$ and $k_{u,t}^*$ by definition. Also, it is easy to see that $k_{l,t}^* < k_{u,t}^*$. If $d_t < k_{l,t}^*$ ($d_t > k_{u,t}^*$), $f_{l,t}(k_v, d_t)$ ($f_{u,t}(k_v, d_t)$) has its minimizer active $k_{l,t}^*$. On the other hand, when $k_{l,t}^* \leq d_t < k_{u,t}^*$, the minimizer is the intersection of the two functions, $k_t^* = d_t$. $\square$

Proof of Proposition 6. The system of equations in Proposition 6 is obtained by taking expectations of Equation (2.16) and using the fact that, at each period, when $d_t = \hat{d}_t$ we are indifferent signing or not signing the contract. $\square$

## A.2. Case Study: The Value of Partial Demand Information

In this section we analyze the value of incorporating partial demand visibility in the contracting decisions when the demand profile is similar to the demand faced by the furniture company. For this purpose, we used a Gamma distribution with periodic shape and scale parameters such that, at each period, its mean and variance equal the empirical mean and variance faced by the company at that specific period. We analyze the five main lanes. The parameters used for the simulation are: $c/c' = 0.3$ (the cost of the spot market with respect to the contract option for the company ranges from 0.25-0.35), $\gamma = 1$, $\tau = 4$ and $\delta = 0.95$. Appendix A4 includes more details about the numerical experiments. First, we note that qualitatively the threshold change from one period to another so the free of commitment periods coordinate with the demand surges.

For each lane, we calculate the value function at each period when the contracting decisions take into account the demand visibility ($\mathbb{V}_{partial,t}$) and the value function when such information is ignored ($\mathbb{V}_{null,t}$). Denoting the sum of the value functions across the season as $\mathbb{V}_{null} = \sum_{t \in T} \mathbb{V}_{null,t}$ and $\mathbb{V}_{partial} = \sum_{t \in T} \mathbb{V}_{partial,t}$, we measure the improvement as $Impro = 100\text{x}(1 - \frac{\mathbb{V}_{partial}}{\mathbb{V}_{null}})$. We show the results in Figure A.1-(a). We see that, in general, considering demand visibility can reduce costs by 3-4% when the commitment horizon is four periods and the company has one period of demand visibility. We also show in Figure A.1-(b) the average seasonal transportation needs of each lane (in normalized units). Note that Lane 1 has the largest volume of transportation needs. It is also the lane that has the lowest Coefficient of Variance, which translates to lower improvement when demand visibility is incorporated. Similarly, Lanes 5-10 have the largest Coefficient of Variance and, therefore, partial demand visibility provides larger cost improvements.

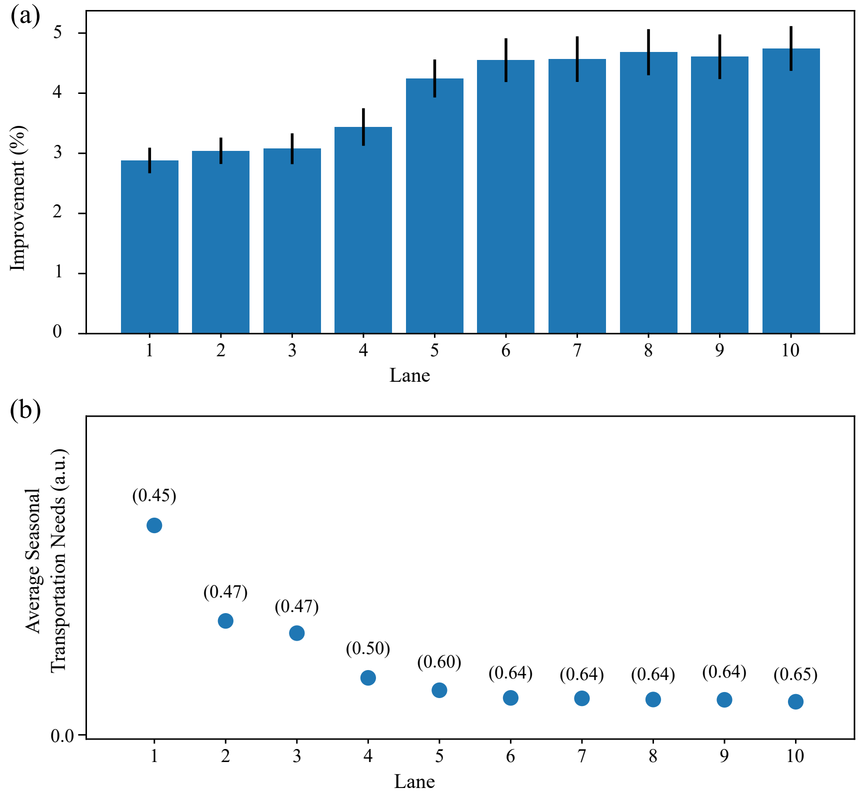


Figure A.1. Results for the 10 main lanes of the furniture company. (a) Cost Reduction when partial demand information is used. (b) Volume of the demand needs of each lane (in arbitrary units). Weekly CV shown in parenthesis.

## A.3. Commitment to a Contract from Among a Discrete set of Capacities

It is straightforward to extend the analysis of single capacity choices to a discrete choice of $M$ capacities $\mathcal{M} = \{k_i, i = 1, 2, ..., M\}$. We assume that $k_i > k_j$ if $i > j$. Let $C^i(d)$ be the cost of serving $d$ units of demand if the decision maker commits to capacity $k_i$. Thus, $C^i(d) = ck_i + c'[d - k_i]^+$. $C^0(d)$ is the cost of serving the entire demand $d$ from the spot market. Figure A.2-(a) shows how cost changes for different levels of capacity commitment. The lighter is the line, the larger is the capacity of the contract.

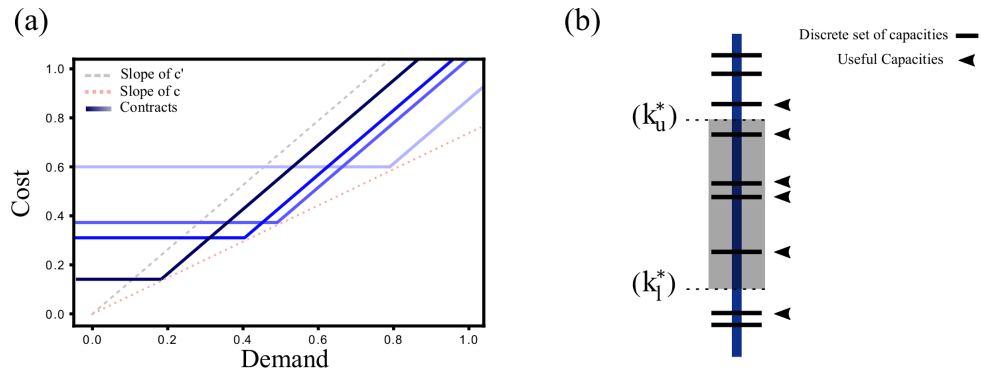


Figure A.2. (a): Cost of four different contracts. (b): Discrete set of capacities vs. set of useful capacities.

Let $V(d)$ be the value function when the next period demand is $d$ and let $\mathbb{V}$ be its the expected value. We denote the expected one period cost when capacity $k_i$ is used as $\beta_i = \mathbb{E}_{D^w}[C^i(D^w)]$. The value function can be expressed as the minimum of: (0) Not signing any contract, (1) signing contract with capacity $k_1$, (2) signing contract with capacity $k_2$, etc. It can be formulated as follows:

(A.4)

$$\begin{aligned} V(d) &= \min\{C^0(d) + \delta\,\mathbb{V}, C^1(d) + \Delta(\tau)\beta_1 + \delta^\tau\,\mathbb{V}, \ldots, C^M(d) + \Delta(\tau)\beta_M + \delta^\tau\,\mathbb{V})\} = \\ &= \min\{C^0(d) + \delta\,\mathbb{V}, C^{M,\tau}(d) + \delta^\tau\,\mathbb{V})\}, \end{aligned}$$

where $C^{M,\tau}(d)$ is the total cost incurred during the commitment horizon $\tau$ when the most cost-efficient capacity is selected over the set $\mathcal{M}$ given observed next period demand $d$:

$$C^{M,\tau}(d) \quad = \min_{i\in\mathcal{M}}(C^i(d) + \Delta(\tau)\beta_i)).$$

The reformulation in the second equality can be interpreted as a trade-off between serving the next period demand entirely from the spot market and committing to a contract.

Assume that it is known that the decision maker will commit to one of the capacities in the set $UC$. In that case, the structure of the solution is simple as all the contract options have the same commitment horizon. The results of analyzing this scenario is close to the continuum set of contracts discussed in Section 2.3.3. It can be shown that the decision maker will commit to one of the capacities in the set $UC$ where: $UC = \{k \in \mathcal{M} | k \in [k_l, k_u]\} \cup \{\min_{\substack{k \in \mathcal{M} \\ k > k_u}} k\} \cup \{\max_{\substack{k \in \mathcal{M} \\ k < k_l}} k\}$. We show in Figure A.2-(b) an example of the $UC$ obtained with a discrete set of capacities. Once this set $UC$ is defined, it is easy to see that the capacity selected will either be the largest capacity in $UC$ that is smaller than visible next period demand d or the smallest capacity in $UC$ that is larger than $d$. One can show that there exists a Discrete-Cost Commitment Balance Condition that determines whether one must ever commit to a capacity contract.

## A.4. Details of the Numerical Experiments in Sections 2.3.4 and the Case of Study (appendix A.2).

The values computed in the numerical experiments of Section 2.3.4 and the case of study are obtained by sampling 75 realizations from the demand distribution and computing the numerical threshold and value function using the value iteration algorithm. We stopped the value function iteration when the difference in the value function and the thresholds where less than 1E-05 from iteration to the other. Each experiment is replicated 10 times. The threshold lines and improvement bars shown are the expected value of the 10 replications and the error bars show the standard error of the computed expectations.

The distribution used for the experiments is Gamma[1] with shape parameter $a$ and scale parameter $\theta$. In order to set an experiment with mean $\mu$ and variance $\sigma$, we set $\theta = \frac{\sigma^2}{\mu}$ and $a = \frac{\mu}{\theta}$. We chose to use Gamma distribution in our experiments because of its flexibility and because, in our assumption, demand is a continuous non-negative random variable. Alternatively, we could have used the Normal distribution for the cases where the Coefficient of Variance is relatively low (e.g. $< 33$). However, from our observation of the empirical data, we see that the Coefficient of Variance of the demand faced by the company at each lane is $0.4 - 0.7$, which makes the probability of negative demand too high. In the next section of the appendix, we repeat the experiments using the absolute Normal distribution. We see that the results do not change.

[1]See Figure A.8-A.10 in appendix A5 for further experiments with an absolute Normal distribution.

## A.5. Selected Supplementary Experiments

We additionally investigated how demand variance and distribution specification affect the results found in Sections 2.3.2 - 2.3.4. We present the results below. The captions summarize the information and analysis of each Figure.

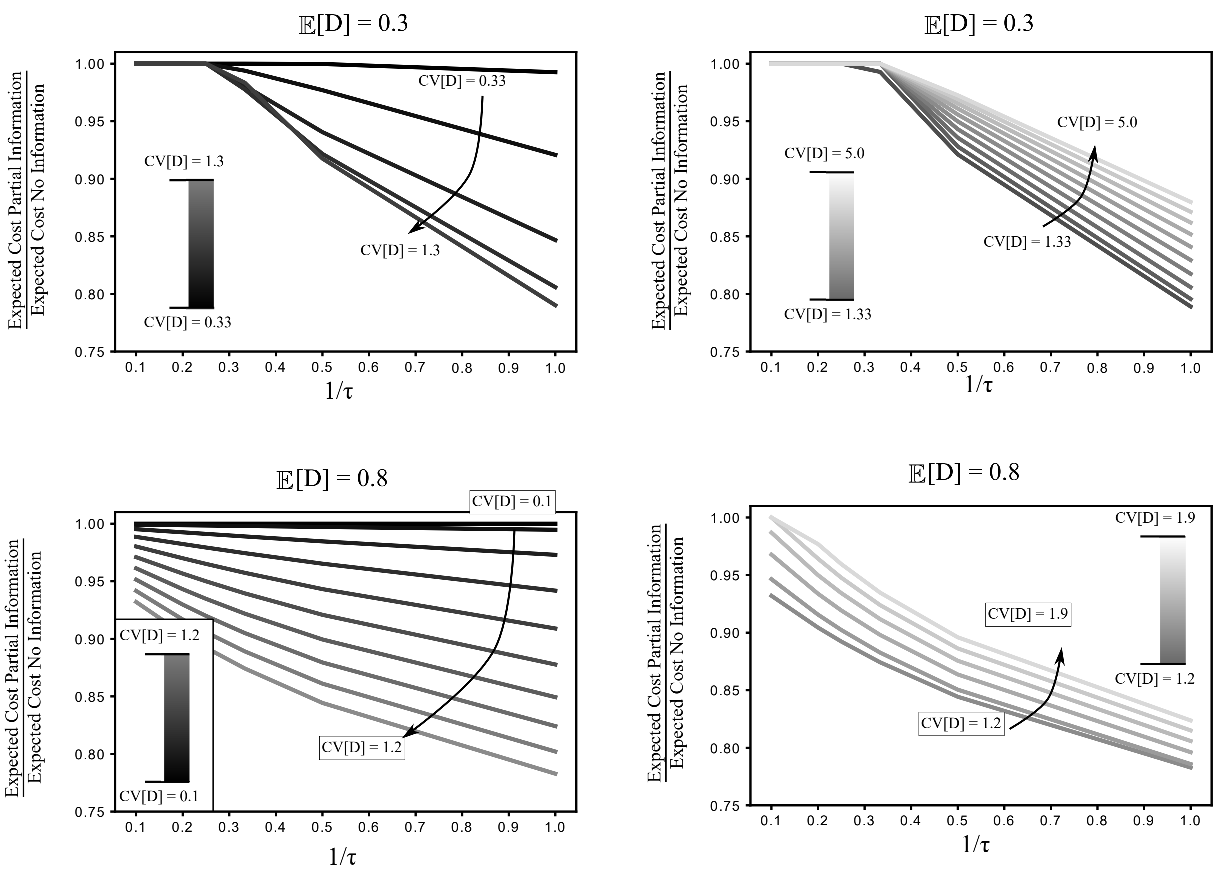


Figure A.3. The value of information when the CV is changed.

Figure A.3 shows the value of information when the CV is changed for a capacity fixed to 1 with both low demand expectation (top) and moderate demand expectation (bottom). Figures on the left show the range of moderate CV and those on the right show high CV. In general terms, as CV increases, the value of information increases as well. When CV reaches extreme values, vale of information decreases. This is because the demand is so variable that using or not using the single capacity does not make a difference.

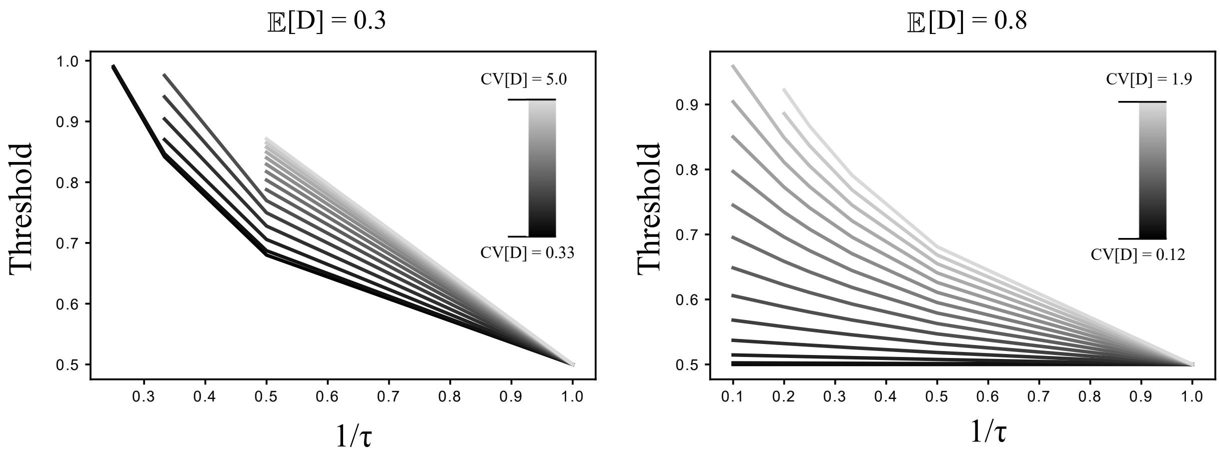


Figure A.4. The threshold with a single capacity when the CV is changed.

Figure A.4 shows the threshold with a single capacity when the CV is changed, for a capacity fixed to 1 with both low demand expectation (left) and moderate demand expectation (right). When the expectation is fixed but the CV increases, demand is with high probability low (or moderate). However, the probability of extremely high occurrences is not negligible (due to the heavy tail). Only when we observe any these large deviation the contract is signed.

Figure A.5 shows the shape of the threshold $\hat{d}$ as a function of expected demand and $1/\tau$, when variance is fixed. For these experiments, we fixed the variance of the demand to 8 and the parameters $k = 80$, $\delta = 0.9$ and $c/c' = 0.5$. Similar shape is obtained in Figure 2.4.

In Figure A.6 we show the value of information when capacity is a decision variable. In these experiments we use a Gamma distribution $\theta = \sigma^2$ and $k = 1/\theta$. Hence, we

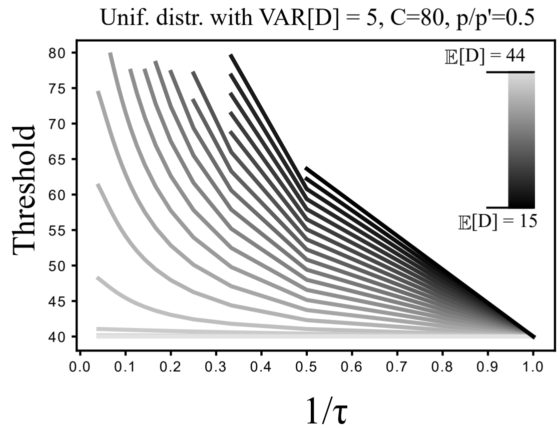


Figure A.5. Fixed Capacity Threshold with Uniform Distribution.

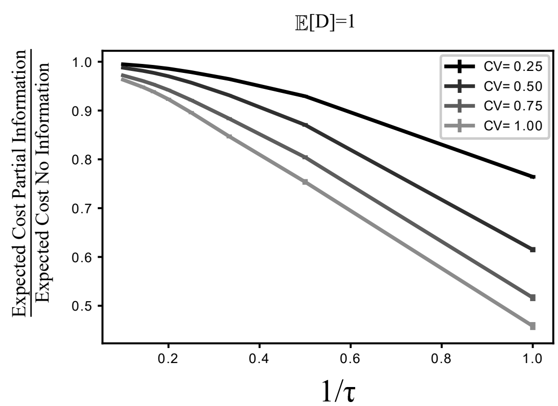


Figure A.6. Value of information when capacity is a decision variable.

fixed mean to 1 and we changed the CV. For these experiments we use the value function iteration algorithm. Demand of each experiment is sampled 50 times. We replicate the experiments 10 times and report the expectation and the standard error. Costs are set $c/c' = 0.3$. The value of partial demand visibility increases with CV.

Figure A.7 we show the threshold when capacity is a decision variable. In these experiments we use a Gamma distribution $\theta = \sigma^2$ and $k = 1/\theta$. Hence, we fix mean to 1 and we change the CV. For these experiments we use the value function iteration algorithm. Demand of each experiment is sampled 50 times. We replicate the experiments

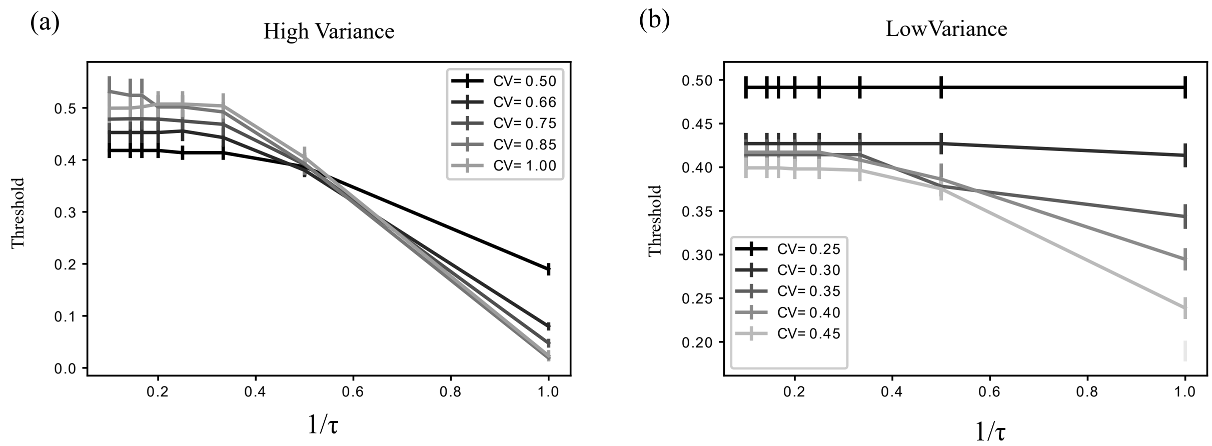


Figure A.7. Threshold when capacity is a decision variable.

10 times and report the expectation and the standard error. Costs are set $c/c' = 0.3$. (a) CV ranges from 1.0-0.5. (b) CV ranges from 0.45-0.25. The threshold is low when the commitment horizon is long as obtained in Figure 2.7.

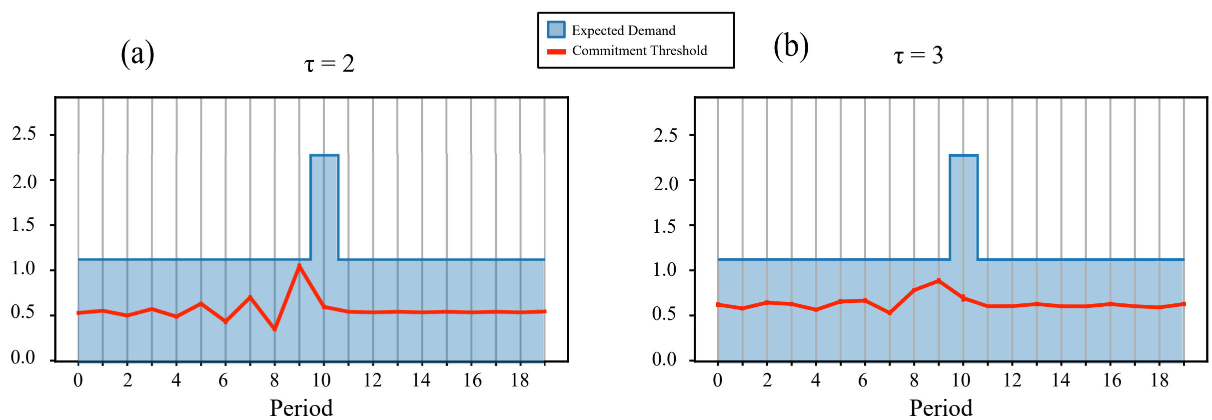


Figure A.8. Same as Figure 2.9 but using an absolute Normal distribution.

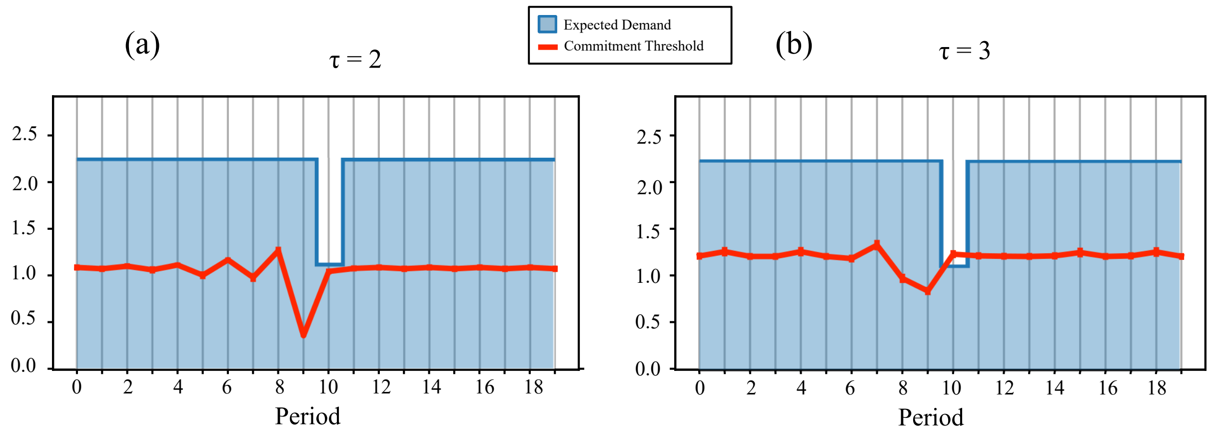


Figure A.9. Same as Figure 2.9 but using an absolute Normal distribution.

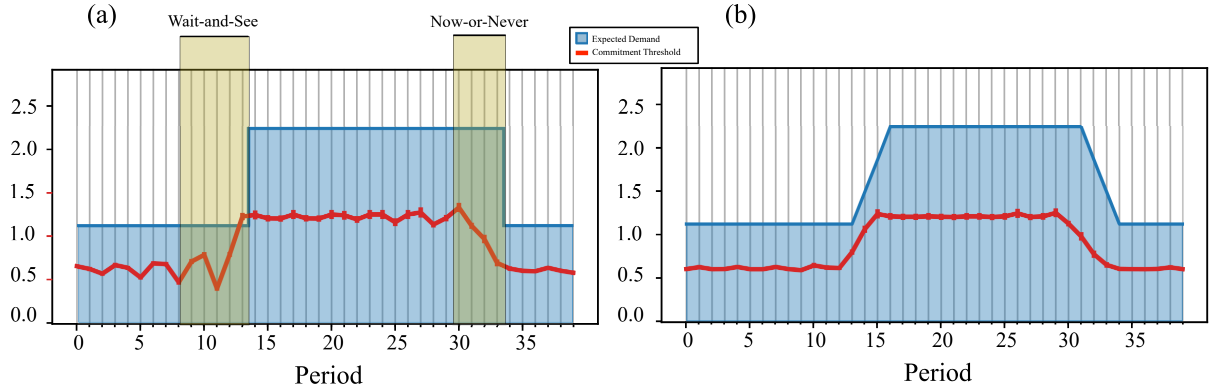


Figure A.10. Same as Figure 2.10 but using an absolute Normal distribution.

APPENDIX B

# Appendix from Chapter 3

## B.1. Proofs

PROOF OF PROPOSITION 7. It is a particular case of Theorem 2. $\square$

PROOF OF THEOREM 2. We first consider the null information case. By assumption $\mathbf{k}^*_{null} \succ 0$, the FOCs of the null information problem imply that $\boldsymbol{\nu} = 0$. Hence, $\mathbf{k}^*_{null}$ is such that it solves:

$$\begin{pmatrix} 0 \\ c_s \\ c_s \end{pmatrix} P(\Omega_I(\mathbf{k})) + \begin{pmatrix} c_s \\ c_s \\ c_s \end{pmatrix} P(\Omega_{III}(\mathbf{k})) + \begin{pmatrix} c_s \\ 0 \\ c_s \end{pmatrix} P(\Omega_{II}(\mathbf{k})) - c = 0. \tag{B.1}$$

Now, consider the partial information problem of minimizing $\mathrm{Cost}_{Total}(\mathbf{k}, \mathbf{d})$ (eq. (B.5)) with respect to $\mathbf{k}$. First we explain the key idea of the proof and then we formally prove the theorem. When the observed demand $\mathbf{d}$ is low enough, the demand is entirely satisfied and the spot market is not used in the first period: $\mathrm{Cost}_{Total}(\mathbf{k}, \mathbf{d}) = \mathbf{c}'\mathbf{k} + (\tau - 1)V(\mathbf{k})$. This problem is continuous and differentiable and can be solved as a regular newsvendor network with costs $\hat{\mathbf{c}} = \mathbf{c}(1 + \frac{1}{\tau-1})$. The solution of this problem is $\boldsymbol{\psi}^0$. With this, we can characterize $\Omega_0(\boldsymbol{\psi}^0)$. Following this idea, we formalize the proof of Theorem 2.

Define $\mathbf{v}^0 = (0,0,0)'$, $\mathbf{v}^I = (0,1,1)$, $\mathbf{v}^{II} = (1,0,1)$ and $\mathbf{v}^{III} = (1,1,1)$ and the $\boldsymbol{\psi}^i$ boundary capacities for $i \in \{0, I, II, III\}$:

(B.2)

$$\boldsymbol{\psi}^i = \{\mathbf{k} \in \mathbb{R}^3_+ \,|\exists \boldsymbol{\nu} \in \mathbb{R}^3_+ \text{ s.t. } 0 = \mathbf{c} - \boldsymbol{\nu} - c^s\mathbf{v}^i + (\tau - 1)\nabla_{\tilde{\mathbf{k}}=\mathbf{k}} V(\tilde{\mathbf{k}}) \text{ and } \boldsymbol{\nu}'\mathbf{k} = 0\}.$$

We prove Theorem 2 using the following Lemma.

**Lemma 4.** *If* $\mathbf{d}^i$ *is an interior point of* $\Omega_i(\boldsymbol{\psi}^i)$ *(*$i = \{0, I, II, III\}$*), then:*

*(a) There exist a neighbourhood around* $\boldsymbol{\psi}^i$ *such that* $Cost_{Total}(\mathbf{k}, \mathbf{d}^i)$ *is smooth.*

*(b)* $\boldsymbol{\psi}^i$ *is a global minimizer.*

**Proof.** Consider the function of the total costs when demand is an interior point of $\Omega_i(\boldsymbol{\psi}^i)$ evaluated at $\boldsymbol{\psi}^i$ ( i.e. $\text{Cost}_{Total}(\boldsymbol{\psi}^i, \mathbf{d}^i)$).

(a) The non-smooth component of the function is the minimization problem of the visible costs ($\pi(\mathbf{k}, \mathbf{d})$). Let us consider only this component and fix $\boldsymbol{\psi}^i$: This defines a set of regions $\Omega_j(\boldsymbol{\psi}^i)$ for $j = \{0, I, II, III\}$ described in Section 3.4.3. Also, as $\mathbf{d}^i$ is an interior point $\Omega_i(\boldsymbol{\psi}^i)$, the dual variables of the minimization are constant [**128**]. Because $\mathbf{d}^i$ is an interior point, there exist a ball around $\boldsymbol{\psi}^i$ such that the set of active constraints does not change and, hence, the dual variables remain constant, which equals the gradient of $\pi(\mathbf{k}, \mathbf{d}^i)$ with respect to $\mathbf{k}$. Therefore, in this neighbourhood, $\text{Cost}_{Total}(\mathbf{k}, \mathbf{d}^i)$ is a sum of smooth functions.

(b) As the function is smooth (within a neighbourhood of $\boldsymbol{\psi}^i$), we can evaluate the gradient of the function at that point:

(B.3)

$$\nabla_{\tilde{\mathbf{k}}=\mathbf{k}}\text{Cost}_{Total}(\tilde{\mathbf{k}},\mathbf{d}^i)) = \mathbf{c} - c^s\mathbf{v}^i + (\tau-1)\nabla_{\tilde{\mathbf{k}}=\mathbf{k}}V(\tilde{\mathbf{k}}) = (\tau-1)(\nabla_{\tilde{\mathbf{k}}=\mathbf{k}}\mathbb{E}_{\mathbf{D}}\,\pi(\tilde{\mathbf{k}},\mathbf{D}) + \hat{\mathbf{c}}^i).$$

with $\hat{\mathbf{c}}^i = \frac{\tau\mathbf{c}-c_s\mathbf{v}^i}{\tau-1} \succ 0$. Now, consider the one-period null-information newsvendor network problem identical to the one defined in Section 3.1 but with capacity prices $\hat{\mathbf{c}}^i$ instead of $\mathbf{c}$. We call the total cost of this problem $\hat{V}(\mathbf{k}) = \mathbb{E}_{\mathbf{D}}\,\pi(\mathbf{k},\mathbf{D}) + \hat{\mathbf{c}}^{i\prime}\mathbf{k}$. As $\hat{\mathbf{c}}^i \succ 0$, $\hat{V}(\mathbf{k})$ is smooth and the gradient exists. Note that $\nabla_{\tilde{\mathbf{k}}=\boldsymbol{\psi}^i}\text{Cost}_{Total}(\tilde{\mathbf{k}},\mathbf{d}_i) = (\tau-1)\nabla_{\tilde{\mathbf{k}}=\boldsymbol{\psi}^i}\hat{V}(\tilde{\mathbf{k}})$. Also, we know that there exist a unique optimal investment vector that minimizes $\hat{V}(\mathbf{k})$. By definition in Eq. (B.2), this vector is $\boldsymbol{\psi}^i$. As $\boldsymbol{d}^i$ is an interior point of $\Omega_i(\boldsymbol{\psi}^i)$, it implies that $\boldsymbol{\psi}^i \succ 0$ (otherwise, any ball around $\boldsymbol{\psi}^i$ would include points outside $\Omega_i(\boldsymbol{\psi}^i)$). Therefore, the Lagrangian multipliers of the KKT for the null information problem defined by $\hat{V}(k)$ are all 0. Hence, $0 = \nabla_{\tilde{\mathbf{k}}=\boldsymbol{\psi}^i}\hat{V}(\tilde{\mathbf{k}}) = \nabla_{\tilde{\mathbf{k}}=\boldsymbol{\psi}^i}\text{Cost}_{Total}(\tilde{\mathbf{k}},\mathbf{d}^i)$. Also, let $\mathbb{H}$ be the Hessian operator, there exist a sufficiently small $\epsilon > 0$ such that for all $v \in \mathbb{R}^3$ and $0 \le \xi \le \epsilon$:

$$\begin{aligned}&\text{Cost}_{Total}(\boldsymbol{\psi}^i + \epsilon\mathbf{v},\mathbf{d}^i) = \text{Cost}_{Total}(\boldsymbol{\psi}^i,\mathbf{d}^i) + \tfrac{1}{2}\mathbf{v}'\mathbb{H}_{\mathbf{k}=\boldsymbol{\psi}^i+\xi\mathbf{v}}\text{Cost}_{Total}(\mathbf{k},\mathbf{d}^i)\mathbf{v} = \\ &\text{Cost}_{Total}(\boldsymbol{\psi}^i,\mathbf{d}^i) + \tfrac{(\tau-1)}{2}\mathbf{v}'\mathbb{H}_{\mathbf{k}=\boldsymbol{\psi}^i+\xi\mathbf{v}}\hat{V}(\mathbf{k})\mathbf{v} > \text{Cost}_{Total}(\boldsymbol{\psi}^i,\mathbf{d}^i).\end{aligned}$$

The first equality uses the Taylor theorem, the second uses equation (B.3) and the second inequality uses the fact that the null information problem is strictly convex.

Therefore, $\boldsymbol{\psi}^i$ is a local minimizer of $\text{Cost}_{Total}(\mathbf{k},\mathbf{d}^i)$. Because the function $\text{Cost}_{Total}(\mathbf{k},\mathbf{d}^i)$ is convex, it is a global minimizer. $\square$

Then, whenever $\mathbf{d}$ is an interior point of $\Omega_i(\boldsymbol{\psi}^i)$, the optimal investment vector is $\boldsymbol{\psi}^i$. We now show that $\Omega_i^* \equiv \Omega_i(\boldsymbol{\psi}^i) \subset \Omega_i(\mathbf{k}^*_{null})$ with the following lemma.

**Lemma 5.** $\Omega_i^* \equiv \Omega_i(\boldsymbol{\psi}^i) \subset \Omega_i(\mathbf{k}^*_{null})$.

**Proof.** We analyze how the each component of $\mathbf{k}^*_{null}$ changes along the path connecting $\mathbf{k}^*_{null}$ and $\boldsymbol{\psi}^i$. From Proposition 3 in [**128**] we have that the sensitivity on the costs of the optimal investment vector under the null information is:

$$\nabla_c \mathbf{k}^*_{null} = \begin{pmatrix} -(\alpha_1+\alpha_2+\alpha_4) & -\alpha_1 & \alpha_1+\alpha_2 \\ -\alpha_1 & -(\alpha_1+\alpha_3+\alpha_5) & \alpha_1+\alpha_3 \\ \alpha_1+\alpha_2 & \alpha_1+\alpha_3 & -(\alpha_1+\alpha_2+\alpha_3+\alpha_6) \end{pmatrix}, \tag{B.4}$$

where $\alpha_j$'s are functions of $\mathbf{k}^*_{null}$. In addition, under conditions of Theorem 2, $\alpha_j > 0$ for all $j = 1, \ldots, 6$.

Consider the capacity contracted for the flexible contract under null information as a function of the contract costs $\mathbf{k}^*_{null;3}(\mathbf{c})$. Now, consider the path in $\mathbb{R}^3$ connecting $\mathbf{k}^*_{null}(\mathbf{c})$ and $\boldsymbol{\psi}^i$:

$$\mathbf{p}(\lambda) = \mathbf{c} + \lambda(\hat{\mathbf{c}}_i - \mathbf{c}).$$

Note that $\mathbf{k}^*_{null}(\mathbf{p}(0)) = \mathbf{k}^*_{null}$ and $\mathbf{k}^*_{null}(\mathbf{p}(1)) = \boldsymbol{\psi}^i$. Now, we analyze the change of the invested capacity component $k^*_{null,j}$ along the path $\mathbf{p}$ at the point $\lambda = 0$ (i.e. the direction derivative with respect to path $\mathbf{p}$):

$$\nabla_{\mathbf{p}|\lambda=0} k^*_{null,j} = \lim_{\lambda\to 0} \frac{k^*_{null,j}(\mathbf{c}+\lambda(\hat{\mathbf{c}}^i-\mathbf{c}))-k^*_{null,j}(\mathbf{c})}{\lambda} = \frac{\partial k^*_{null,j}}{\partial p_1}\frac{dp_1}{d\lambda} + \frac{\partial k^*_{null,j}}{\partial p_2}\frac{dp_2}{d\lambda} + \frac{\partial k^*_{null,j}}{\partial p_3}\frac{dp_3}{d\lambda},$$

where

$$\frac{dp_j}{d\lambda} = \frac{1}{\tau-1}(c_j - c_s v^i_j).$$

We now separate the analysis for each region:

$\hat{\Omega}_0$: As argued above, by construction, $\boldsymbol{\psi}^0$ is such that it is the minimizer of a null information single period problem with $\hat{\boldsymbol{c}}^0 = \mathbf{c} + \frac{1}{\tau-1}\mathbf{c} \succ \mathbf{c}$. Because the new newsvendor network has increased all three costs, intuitively the total investment in capacity is reduced. We show this intuition formally as similar steps will be used to prove Theorem 4.

First, we build the path $\mathbf{p}(\lambda) = \mathbf{c} + \lambda(\hat{\mathbf{c}}^0 - \mathbf{c})$ connecting $\boldsymbol{\psi}^0$ and $\mathbf{k}^*_{null}$. Define $A(\mathbf{p}(\lambda)) = \mathbf{k}^*_{null,1}(\mathbf{p}(\lambda)) + \mathbf{k}^*_{null,2}(\mathbf{p}(\lambda)) + \mathbf{k}^*_{null,3}(\mathbf{p}(\lambda))$, i.e. the total investment in capacity. Note that in $\Omega^*_0$ we have that $\frac{dp_j}{d\lambda} = \frac{c_j}{\tau-1}$. Then, we can compute the derivative of $A$ along the path:

$$\nabla_{\mathbf{p}|\lambda=0} A = \sum_{i=1,2,3} \nabla_{\mathbf{p}|\lambda=0} k^*_{null,j} = -(\alpha_1+\alpha_4)\frac{dp_1}{d\lambda} - (\alpha_1+\alpha_5)\frac{dp_2}{d\lambda} + (\alpha_1-\alpha_6)\frac{dp_3}{d\lambda} =$$
$$\frac{-1}{\tau-1}(\alpha_4 c_1 + \alpha_5 c_2 + \alpha_6 c_3 + \alpha_1(c_1+c_2-c_3)) < 0.$$

The last inequality is true since $c_1+c_2 > c_3$. Now, define $B(\mathbf{p}(\lambda)) = k^*_{null,1}(\mathbf{p}(\lambda))+ k^*_{null,3}(\mathbf{p}(\lambda))$. We evaluate how this change along the path:

$$\nabla_{\mathbf{p}|\lambda=0} B = -\alpha_4 \tfrac{dp_1}{d\lambda} - (\alpha_3+\alpha_6)\tfrac{dp_3}{d\lambda} < 0.$$

Finally, we obtain that for $C(\mathbf{p}(\lambda)) = k^*_{null,2}(\mathbf{p}(\lambda))+k^*_{null,3}(\mathbf{p}(\lambda))$ that $\nabla_{\mathbf{p}|\lambda=0} C < 0$.

Hence, we have obtained that $\psi^0_1+\psi^0_2+\psi^0_3 < k^*_{null;1}+k^*_{null;2}+k^*_{null;3}$, $\psi^0_1+\psi^0_3 < k^*_{null;1}+k^*_{null;3}$ and $\psi^0_2+\psi^0_3 < k^*_{null;2}+k^*_{null;3}$. Therefore, $\hat{\Omega}_0 \subset \Omega_0(\mathbf{k}^*_{null})$.

$\hat{\Omega}_I$: $\boldsymbol{\psi}^I$ is such that it is the minimizer of a null information single period problem with:

$$\hat{\mathbf{c}}^I = \mathbf{c} + \tfrac{1}{\tau-1}(\mathbf{c} - c_s\mathbf{v}^I) = \begin{pmatrix} c_1(1+\frac{1}{\tau-1}) \\ c_2 - \frac{1}{\tau-1}(c_s-c_2) \\ c_3 - \frac{1}{\tau-1}(c_s-c_3) \end{pmatrix}.$$

Building the path connecting $\boldsymbol{\psi}^I$ and $k^*_{null}$ and following the same steps as for $\Omega_0$: Consider $A = k^*_{null,1}(\mathbf{p}(\lambda))$.

$$\nabla_{\mathbf{p}|\lambda=0} A = -(\alpha_1+\alpha_2+\alpha_4)(\tfrac{c_1}{\tau-1}) + \alpha_1(\tfrac{c_s-c_2}{\tau-1}) - (\alpha_1+\alpha_2)(\tfrac{c_s-c_3}{\tau-1}) < 0.$$

Now, $B = k^*_{null,2}(\mathbf{p}(\lambda)) + k^*_{null,3}(\mathbf{p}(\lambda))$:

$$\nabla_{\mathbf{p}|\lambda=0} B = +\alpha_2(\tfrac{c_1}{\tau-1}) + \alpha_5(\tfrac{c_s-c_2}{\tau-1}) + (\alpha_2+\alpha_6)(\tfrac{c_s-c_3}{\tau-1}) > 0$$

The last inequality is a consequence of the assumption $\mathbf{c} \succ \frac{c_s}{\tau}$. Hence, the optimal investment vector of this problem is such that $\psi_1^I \geq k^*_{null,1}$, $\psi_2^I + \psi_3^I \geq k^*_{null,2} + k^*_{null,3}$. (capacity cost of contract 1 is larger and cost of contract 2 and 3 are lower.) Therefore, $\hat{\Omega}_I \subset \Omega_I(\mathbf{k}^*_{null})$.

$\hat{\Omega}_{III}$: $\boldsymbol{\psi}^{III}$ is such that it is the minimizer of a null information single period problem with:

$$\hat{\mathbf{c}}^{III} = \mathbf{c} + \tfrac{1}{\tau-1}(\mathbf{c} - c_s\mathbf{v}^{III}) = \begin{pmatrix} c_1 - \frac{1}{\tau-1}(c_s - c_1) \\ c_2 - \frac{1}{\tau-1}(c_s - c_2) \\ c_3 - \frac{1}{\tau-1}(c_s - c_3) \end{pmatrix}.$$

Therefore, $\hat{\Omega}_{III} \subset \Omega_{III}(\mathbf{k}^*_{null})$

$\hat{\Omega}_{II}$: With the same arguments we obtain that $\hat{\Omega}_{II} \subset \Omega_{II}(\mathbf{k}^*_{null})$

□

Finally, because the regions $\Omega_i(\mathbf{k}^*_{null})$ are disjoint, the regions $\hat{\Omega}_i$ are disjoint. Also, because $\Omega_i(\mathbf{k})$ are convex for any $\mathbf{k}$, in particular they are convex for $\Omega_i(\boldsymbol{\psi}^i)$. □

Proof of Theorem 3. The first bullet is a direct consequence of Theorem 2. For proving the second bullet, we need to show that the contracted capacity is always a convex combination of the $\boldsymbol{\psi}^i$'s.

For a given observed demand vector $\mathbf{d}$, the total costs can be written as follows:

$$\min_{\mathbf{k}\in\mathbb{R}^3_+} \text{Cost}_{Total}(\mathbf{k},\mathbf{d}) = \min_{\mathbf{k}\in\mathbb{R}^3_+} \pi(\mathbf{k},\mathbf{d}) + \mathbf{c}'\mathbf{k} + (\tau-1)V(\mathbf{k}) =$$

$$= \min_{\mathbf{k}\in\mathbb{R}^3_+} \min_{\substack{\mathbf{y}\in\mathbb{R}^4_+\\ y_1+y_3\le d_1\\ y_2+y_4\le d_2\\ y_1\le k_1\\ y_2\le k_2\\ y_3+y_4\le k_3}} c_s(d_1-y_1-y_3) + c_s(d_2-y_2-y_4) + \mathbf{c}'\mathbf{k} + (\tau-1)V(\mathbf{k})$$

$$\min_{\mathbf{k}\in\mathbb{R}^3_+} \max_{i\in\{0,I,II,III\}} \{(\boldsymbol{\mu}_i^*)'\mathbf{d} + (\boldsymbol{\lambda}_i^* + \boldsymbol{c})'\mathbf{k} + (\tau-1)V(\mathbf{k})\} = \min_{\mathbf{k}\in\mathbb{R}^3_+} \max_{i\in\{0,I,II,III\}} \{\phi_i(\mathbf{k},\mathbf{d}))\}.$$

Where $\boldsymbol{\mu}_i^*$ and $\boldsymbol{\lambda}_i^*$ are the set of extreme points of the feasible set of the dual variables of constraints (2) and (3) respectively. Note that each $\phi_i(\mathbf{k},\mathbf{d})(= (\boldsymbol{\mu}_i^*)'\mathbf{d} + (\boldsymbol{\lambda}_i^* + c)'\mathbf{k} + (\tau - 1)V(\mathbf{k}))$ is a strongly convex function whose minimizer is $\boldsymbol{\psi}^i$. The following two Lemmas prove the Theorem.

**Lemma 6.** *Let $\mathbf{x}_i^*$ for $i \in \mathcal{I} = 1, 2, \ldots, I$ a countable and finite set of points in $\mathbb{R}^n$. Then, if $x \notin CH(\{\mathbf{x}_i^*|i \in \mathcal{I}\})$ (Convex Hull), there exist a direction $\mathbf{v}$ such that for a small enough $\epsilon > 0$, $||\mathbf{x} + \epsilon\mathbf{v} - \mathbf{x}_i||^2 < ||\mathbf{x} - \mathbf{x}_i||^2$ for all $i \in \mathcal{I}$.*

**Proof.** Consider the closest facet of the convex hull with respect to $x$. Now, consider the hyperplane $h(x, \{\mathbf{x}_i^*|i \in \mathcal{I}\})$ containing this facet, we simply call it $h$. Now, let $\mathbf{x}_p \in h$ be the orthogonal projection of $x$ on $h$. Any point in the segment $\mathbf{s}(\lambda) = \lambda\mathbf{x}_p + (1-\lambda)\mathbf{x}$ $(0 < \lambda \le 1)$ is closer to all the points in $h$ than $x$, including all points in the facet. Similarly, it is closer than any point in the $CH(\{\mathbf{x}_i^*|i \in \mathcal{I}\})$, including the extreme points. Let $\epsilon = ||\mathbf{x}_p - \mathbf{x}||$ and $\mathbf{v} = \frac{\mathbf{x}_p - \mathbf{x}}{||\mathbf{x}_p - \mathbf{x}||}$. □

**Lemma 7.** *Let $\phi_i(\mathbf{x})$ for $i \in \mathcal{I}$ be a countable and finite set of strongly convex functions $\mathbb{R}^n \to \mathbb{R}$ with minimizers $\mathbf{x}_i^*$. Then, the minimizer of $\Phi(\mathbf{x}) = \max_{i \in \mathcal{I}} \phi_i(\mathbf{x})$ exists and can be expressed as a convex combination of $\mathbf{x}_i^*$.*

**Proof.** We prove the lemma by showing that for any point that is not in the convex hull, there exist a direction that decreases $\Phi(\mathbf{x})$ (by getting closer to all the minimizers of $i \in \mathcal{I}$). Let $\mathbf{x} \notin CH(\mathbf{x}_i^* | i \in \mathcal{I})$. Then from Lemma 6, we know that there exist a direction such that for small enough $\epsilon > 0$ the point $\mathbf{x}' = \mathbf{x} + \epsilon \mathbf{v}$ is closer to *all* the extreme points $\mathbf{x}_i^*$. Because the functions $\phi_i(\mathbf{x})$ are all *strictly* convex, $\phi_i(\mathbf{x}') < \phi_i(\mathbf{x})$ for all $i \in \mathcal{I}$ and small enough $\epsilon$. Hence, $\Phi(\mathbf{x}') < \Phi(\mathbf{x})$ and, therefore $\mathbf{x}$ cannot be a minimizer. □

The total costs is a particular case of the Lemma above. □

Proof of Proposition 9. The results are obtained by comparing $\hat{\mathbf{c}}^i$ and considering the sensitivity of the optimal capacity vector with respect to the costs (Equation B.4). □

Proof of Proposition 10. If $c_1 < \frac{c_s}{\tau}$, then $(\tau - 1)c_1 < c_s - c_1$. Therefore, the marginal cost of not fulfilling a unit of the next period demand is larger than the cost of carrying a unit of capacity for the rest of the commitment horizon. Hence, it is never optimal to have unsatisfied demand. Therefore, the larger is demand 1 the larger is the investment in capacity 1. By definition there is no bounding region $\Omega_I$ or $\Omega_{III}$. The same line of reasoning can be applied for the case $c_2 < \frac{c_s}{\tau}$. □

Proof of Proposition 8. The inequalities are a direct consequence of Theorem 2. □

PROOF OF THEOREM 4. We show that under certain conditions, we can guarantee that the capacity of the flexible contract in $\Omega_I$ and $\Omega_{II}$ is higher than the capacity of the flexible contract of the null information scenario. We follow the same steps as in Theorem 2.

For $\hat{\Omega}_I$: We analyze how $k^*_{null,3}$ change along the path $\mathbf{p}$ connecting $\mathbf{k}^*_{null}$ and $\boldsymbol{\psi}^I$.

$$\nabla_{\mathbf{p}|\lambda=0} k^*_{null,3} = \sum_{i=1,2,3} \frac{\partial k^*_{null,3}(\mathbf{p})}{\partial p_i} \frac{dp_i}{d\lambda} = (\alpha_1 + \alpha_2)\tfrac{c_1}{\tau-1} - (\alpha_1 + \alpha_3)\tfrac{c_s - c_2}{\tau-1} + (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_6)\tfrac{c_s - c_3}{\tau-1}.$$

Assume that $c_3 = c_2(\geq c_1)$. We then obtain that $\nabla_{\mathbf{p}|\lambda=0} k^*_{null,3} > 0$. Therefore, for small enough $c_3$ (but strictly greater than $c_2$) we obtain that $\psi^I_3 > k^*_{null,3}$.

For $\hat{\Omega}_{II}$: We analyze how $k^*_{null,3}$ change along the path $\mathbf{p}$ connecting $\mathbf{k}^*_{null}$ and $\boldsymbol{\psi}^{II}$.

$$\nabla_{\mathbf{p}|\lambda=0} k^*_{null,3} = -(\alpha_1 + \alpha_2)\tfrac{c_s - c_1}{\tau-1} + (\alpha_1 + \alpha_3)\tfrac{c_2}{\tau-1} + (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_6)\tfrac{c_s - c_3}{\tau-1}.$$

If $c_3 = c_1$, $\nabla_{\mathbf{p}|\lambda=0} k^*_{null,3} > 0$. Therefore, for close enough $c_1$ and $c_3$, $\psi^{II}_3 > k^*_{null,3}$.

$\square$

PROOF OF THEOREM 5. From Proposition 2 in [**128**], if $c_3 > \bar{c}_3 \equiv c_1 + c_2 - c_s \mathbb{P}(\Omega_{III}(\bar{\mathbf{k}}))$, then the flexible contract of the null contract is 0 ($\bar{\mathbf{k}}$ is the solution when only dedicated contracts are used). By optimality we have that the expected costs under partial information scenario in the network ($\mathbb{E}_{\mathbf{D}}[\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D}) | c_3 < c_1 + c_2]$) are lower than the expected costs under the partial information scenario without the network ($\mathbb{E}_{\mathbf{D}}[\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D}) | c_3 \geq c_1 + c_2]$). By fixing the cost of the contract at $\bar{c}_3$ we obtain that the $VoI$ is greater or equal than the $VoI_D$. (In both cases the cost of the null scenario is exactly the same as $k_{null,3} = 0$ without the network or with the network with $c_3 \geq \bar{c}_3$.)

It is easy to see that if there exist observations for which $k_3^*(\mathbf{d}) > 0$ at the point $c_3 = \bar{c}_3$, then the inequality holds strictly. While it is a mild condition, in Lemma 8 we provide two sufficient conditions showing that strict inequality can hold. Finally, in Lemma 9 we show that independence is a sufficient condition.

**Lemma 8.** *[**Two Sufficient Conditions for Strict Inequality**] Assume that $f(d_1, d_2) > 0$ for all $d_1, d_2 \geq 0$. There always exists a flexible contract cost $\tilde{c}_3 < c_1 + c_2$ such that $VoI > VoI_{-F}$, any of the conditions below hold:*

*(a)* $(1 + \frac{1}{\tau-1})\mathbb{P}(D_1 > \bar{k}_1, D_2 > \bar{k}_2) \geq \mathbb{P}(D_1 > \psi_1^u, D_2 > \psi_2^l)$,

*(b)* $(1 + \frac{1}{\tau-1})\mathbb{P}(D_1 > \bar{k}_1, D_2 > \bar{k}_2) \geq \mathbb{P}(D_1 > \psi_1^l, D_2 > \psi_2^u)$.

*(Where:* $\bar{k}_i = F_i^{-1}(\frac{c_s - c_i}{c_s})$, $\psi_i^l = F_i^{-1}(\frac{c_s - c_i}{c_s} - \frac{c_i}{(\tau-1)c_s})$ *and* $\psi_i^u = F_i^{-1}(\frac{c_s - c_i}{c_s} + \frac{c_s - c_i}{c_s(\tau-1)})$. *)*

**Proof.** Assume that the first condition holds. Also, assume that the cost of the flexible contract is $\bar{c}_3 = c_1 + c_2 - c_s\mathbb{P}(\Omega_{III}(\bar{\mathbf{k}}))$ and, hence, $k_{3,null}^* = 0$. As similarly argued in the proof of Theorem 2, for large enough $d_1$ and low enough $d_2$, the total costs can be written as follows:

$$\text{Cost}_{Total}(\mathbf{k}, \mathbf{d}) = c_s(d_2 - k_1 - k_3) + \mathbf{c}'\mathbf{k} + (\tau - 1)V(\mathbf{k}).$$

We show that when such a demand realization, $\boldsymbol{d}$, is observed, $k^*(\boldsymbol{d})_3 > 0$, and, therefore, $VoI > VoI_{-F}$. The minimizer of the above equation also minimizes a single-period null information newsvendor network problem with costs $\hat{\mathbf{c}} = (c_1 - \frac{c_s - c_1}{\tau-1}, c_2 + \frac{c_2}{\tau-1}, \bar{c}_3 - \frac{c_s - \bar{c}_3}{\tau-1})$. We show that for this problem, the threshold of the flexible contract cost $\bar{\hat{c}}_3 = \hat{c}_1 + \hat{c}_2 - c_s\mathbb{P}(\Omega_{III}(\bar{\hat{\mathbf{k}}}))$ is such that $\bar{\hat{c}}_3 > \hat{c}_3 = \bar{c}_3 - \frac{c_s - \bar{c}_3}{\tau-1}$, where $\bar{\hat{\mathbf{k}}}$ is the optimal capacity with costs

$\hat{\mathbf{c}}$ when only the dedicated contracts are used.

$$
\begin{aligned}
&\bar{\hat{c}}_3 = \hat{c}_1 + \hat{c}_2 - c_s\mathbb{P}(\Omega_{III}(\bar{\hat{\mathbf{k}}})) = \\
&= c_1 - \tfrac{c_s - c_1}{\tau-1} + c_2 + \tfrac{c_2}{\tau-1} - c_s\mathbb{P}(\Omega_{III}(\bar{\hat{\mathbf{k}}})) + c_s\mathbb{P}(\Omega_{III}(\bar{\mathbf{k}})) - c_s\mathbb{P}(\Omega_{III}(\bar{\mathbf{k}})) = \\
&= \bar{c}_3 + \tfrac{c_1 + c_2 - c_s}{\tau-1} + c_s(\mathbb{P}(\Omega_{III}(\bar{\mathbf{k}})) - \mathbb{P}(\Omega_{III}(\bar{\hat{\mathbf{k}}}))) = \\
&= \hat{c}_3 + \tfrac{c_1 + c_2 - \bar{c}_3}{\tau-1} + c_s(\mathbb{P}(\Omega_{III}(\bar{\mathbf{k}})) - \mathbb{P}(\Omega_{III}(\bar{\hat{\mathbf{k}}}))) = \\
&= \hat{c}_3 + c_s((1 + \tfrac{1}{\tau-1})\mathbb{P}(\Omega_{III}(\bar{\mathbf{k}})) - \mathbb{P}(\Omega_{III}(\bar{\hat{\mathbf{k}}}))).
\end{aligned}
$$

Note that $\bar{\mathbf{k}} = (F_1^{-1}(\frac{c_s - c_1}{c_s}), F_2^{-1}(\frac{c_s - c_2}{c_s}), 0)$ and $\bar{\hat{\mathbf{k}}} = (F_1^{-1}(\frac{c_s - c_1}{c_s} - \frac{c_s - c_1}{c_s(\tau-1)}), F_2^{-1}(\frac{c_s - c_2}{c_s} + \frac{c_2}{c_s(\tau-1)}), 0)$ and $\mathbb{P}(\Omega_{III}(\mathbf{k})) = \mathbb{P}(D_1 > k_1, D_2 > k_2)$. Hence, by the assumption in (a) the last term of the last equality is positive. Hence, $\bar{\hat{c}}_3 > \hat{c}_3$ and, therefore, there exists a region in the sample space such that $k_3^* > 0$, which implies that $VoI > VoI_{-F}$. The same argument can be used to proof the Lemma for condition (b). Finally, note that these two conditions depend only on the demand distribution and costs $c_1$, $c_2$ and $c_s$. □

**Lemma 9.** *[**Independence is a Sufficient Condition**] If $D_1$ and $D_2$ are independent, condition (a) and (b) of Lemma 8 always hold.*

**Proof.** Assume that $D_1$ and $D_2$ are independent, hence $\mathbb{P}(\Omega_{III}(\mathbf{k})) = \mathbb{P}(D_1 > k_1)\mathbb{P}(D_2 > k_2)$. Let $\bar{\hat{\mathbf{k}}} = (F_1^{-1}(\frac{c_s - c_1}{c_s} - \frac{c_s - c_1}{c_s(\tau-1)}), F_2^{-1}((\frac{c_s - c_2}{c_s} + \frac{c_2}{c_s(\tau-1)}), 0)$. With simple algebra we obtain:

$$
c_s((1 + \tfrac{1}{\tau-1})\mathbb{P}(\Omega_{III}(\bar{\mathbf{k}})) - \mathbb{P}(\Omega_{III}(\bar{\hat{\mathbf{k}}}))) > 0 \iff c_s > c_2.
$$

The last is always true by assumption. The same can be shown to prove condition (b). □

□

Proof of Theorem 6. Note that the first term of the AVoI is $V(\mathbf{k}_{null})$, which a single period newsvendor network with null information. [**128**] shows that $\frac{\partial V(\mathbf{k}_{null})}{\partial c_3} = k^*_{3,null}$.

Using similar arguments, the derivative of the second term of AVoI, $\mathbb{E}_{\mathbf{D}}[\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D})]$, is $\frac{\mathbb{E}_{\mathbf{D}}[\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}),\mathbf{D})]}{\partial c_3} = \mathbb{E}_{\mathbf{D}}[k^*_3(\mathbf{D})]$. Note that $\text{Cost}_{Total}(\boldsymbol{k}, \boldsymbol{d})$ is concave with respect to $\boldsymbol{c}$ for all $\boldsymbol{k}$ in the convex set $\mathbb{R}^3_+$. In this case, convexity is preserved under minimization, and, therefore, $\min_{\boldsymbol{k}\in\mathbb{R}^3_+} \text{Cost}_{Total}(\boldsymbol{k}, \boldsymbol{d})$ is concave with respect to $\boldsymbol{c}$. In addition, since $\boldsymbol{D}$ is continuous and $\min_{\boldsymbol{k}\in\mathbb{R}^3_+} \text{Cost}_{Total}(\boldsymbol{k}, \boldsymbol{d})$ is concave in $\boldsymbol{c}$, the partial derivative and the expectation can be interchanged [**67**]. □

Proof of Theorem 7. Theorem 7 is a consequence of Theorem 6 and Theorem 5. □

Proof of Theorem 8. Note that it is optimal to always commit to at least $d_l$ units of dedicated capacity at each node. To ease the formulation, we shift the demand at each node by $d_l$ units. In other words, we consider the case where demand at each node can take two realizations: $d_l = 0$ and $d_h = 2\delta$. The only impact of that is a shift of the costs equal to $c_1 d_l + c_2 d_l$.

We first describe the costs for each capacity vector and demand realization. Under the null information scenario, the cost for each possible commitment vector is:

$\underline{\boldsymbol{k}_a = (0,0,0)'}$: $Cost \equiv Cost_{n,a} = 2\delta\tau c_s(p_{hl} + p_{lh} + 2p_{hh})$.

$\underline{\boldsymbol{k}_b = (0,2\delta,0)'}$: $Cost \equiv Cost_{n,b} = 2\delta\tau(c_1 + c_s(p_{lh} + p_{hh}))$.

$\underline{\boldsymbol{k}_c = (2\delta,0,0)'}$: $Cost \equiv Cost_{n,c} = 2\delta\tau(c_2 + c_s(p_{lh} + p_{hh}))$.

$\underline{\boldsymbol{k}_d = (2\delta,2\delta,0)'}$: $Cost \equiv Cost_{n,d} = 2\delta\tau(c_1 + c_2)$.

$\underline{\boldsymbol{k}_e = (0,0,2\delta)'}$: $Cost \equiv Cost_{n,e} = 2\delta\tau(c_3 + c_s p_{hh})$.

From the previous expression, we can obtain that when $c_3 \geq \bar{c}_3 = \min\{c_1 + c_2 - c_s p_{hh}, c_1 + c_s p_{lh}, c_s(1 - p_{ll})\}$, the decision maker never uses flexible capacity under the null information scenario. Note that for $\bar{c}_3$ to be $\bar{c}_3 = c_1 + c_s p_{lh}$, it needs two conditions that never meet simultaneously.

(a) $c_1 + c_s p_{lh} < c_1 + c_2 - c_s p_{hh}$. When $c_1 = c_2$, this occurs if and only if $c_1 > c_s(p_{hh} + p_{lh})$.

(b) $c_1 + c_s p_{lh} < c_s(1 - p_{ll}) = c_s(p_{hl} + p_{lh} + p_{hh})$. When $p_{hl} = p_{lh}$, this occurs if and only if $c_1 < c_s(p_{hh} + p_{lh})$.

Therefore, $\bar{c}_3 = \min\{c_1 + c_2 - c_s p_{hh}, c_s(1 - p_{ll})\}$. With partial demand visibility scenario, there exist 4 possible demand realizations in the first period. We below we show the cost for each possible scenario:

- $\boldsymbol{d} = (0, 0)$:

  $\underline{\boldsymbol{k}_a = (0, 0, 0)'}$: $Cost_{ll,a} = 2\delta(\tau - 1)c_s(p_{hl} + p_{lh} + 2p_{hh})$.

  $\underline{\boldsymbol{k}_b = (0, 2\delta, 0)'}$: $Cost_{ll,a} = 2\delta\{(\tau - 1)(c_1 + c_s(p_{lh} + p_{hh})) + c_1\}$.

  $\underline{\boldsymbol{k}_c = (2\delta, 0, 0)'}$: $Cost_{ll,a} = 2\delta\{(\tau - 1)(c_2 + c_s(p_{lh} + p_{hh})) + c_2\}$.

  $\underline{\boldsymbol{k}_d = (2\delta, 2\delta, 0)'}$: $Cost_{ll,a} = 2\delta\{(\tau - 1)(c_1 + c_2) + (c_1 + c_2)\}$.

  $\underline{\boldsymbol{k}_e = (0, 0, 2\delta)'}$: $Cost_{ll,a} = 2\delta\{(\tau - 1)(c_3 + c_s p_{hh}) + c_3\}$.

- $\boldsymbol{d} = (0, 2\delta)$:

  $\underline{\boldsymbol{k}_a = (0, 0, 0)'}$: $Cost_{lh,a} = 2\delta\{(\tau - 1)c_s(p_{hl} + p_{lh} + 2p_{hh}) + c_s\}$.

  $\underline{\boldsymbol{k}_b = (0, 2\delta, 0)'}$: $Cost_{lh,b} = 2\delta\{(\tau - 1)(c_2 + c_s(p_{lh} + p_{hh})) + c_2\}$.

  $\underline{\boldsymbol{k}_c = (2\delta, 0, 0)'}$: $Cost_{lh,c} = 2\delta\{(\tau - 1)(c_1 + c_s(p_{lh} + p_{hh})) + c_1 + c_s\}$.

  $\underline{\boldsymbol{k}_d = (2\delta, 2\delta, 0)'}$: $Cost_{lh,d} = 2\delta\{(\tau - 1)(c_1 + c_2) + (c_1 + c_2)\}$.

  $\underline{\boldsymbol{k}_e = (0, 0, 2\delta)'}$: $Cost_{lh,e} = 2\delta\{(\tau - 1)(c_3 + c_s p_{hh}) + c_3\}$.

- $\boldsymbol{d} = (2\delta, 0)$:

  $\underline{\boldsymbol{k}_a = (0,0,0)'}$: $Cost_{hl,a} = 2\delta\{(\tau-1)c_s(p_{hl}+p_{lh}+2p_{hh})+c_s\}$.

  $\underline{\boldsymbol{k}_b = (0,2\delta,0)'}$: $Cost_{hl,b} = 2\delta\{(\tau-1)(c_1+c_s(p_{lh}+p_{hh}))+c_1+c_s\}$.

  $\underline{\boldsymbol{k}_c = (2\delta,0,0)'}$: $Cost_{hl,c} = 2\delta\{(\tau-1)(c_2+c_s(p_{lh}+p_{hh}))+c_2\}$.

  $\underline{\boldsymbol{k}_d = (2\delta,2\delta,0)'}$: $Cost_{hh,d} = 2\delta\{(\tau-1)(c_1+c_2)+(c_1+c_2)\}$.

  $\underline{\boldsymbol{k}_e = (0,0,2\delta)'}$: $Cost_{hl,e} = 2\delta\{(\tau-1)(c_3+c_sp_{hh})+c_3\}$.

- $\boldsymbol{d} = (2\delta, 2\delta)$

  $\underline{\boldsymbol{k}_a = (0,0,0)'}$: $Cost_{hh,a} = 2\delta\{(\tau-1)c_s(p_{hl}+p_{lh}+2p_{hh})+c_s+c_s\}$.

  $\underline{\boldsymbol{k}_b = (0,2\delta,0)'}$: $Cost_{hh,b} = 2\delta\{(\tau-1)(c_1+c_s(p_{lh}+p_{hh}))+c_1+c_s\}$.

  $\underline{\boldsymbol{k}_c = (2\delta,0,0)'}$: $Cost_{hh,c} = 2\delta\{(\tau-1)(c_2+c_s(p_{lh}+p_{hh}))+c_2+c_s\}$.

  $\underline{\boldsymbol{k}_d = (2\delta,2\delta,0)'}$: $Cost_{hh,d} = 2\delta\{(\tau-1)(c_1+c_2)+(c_1+c_2)\}$.

  $\underline{\boldsymbol{k}_e = (0,0,2\delta)'}$: $Cost_{hh,e} = 2\delta\{(\tau-1)(c_3+c_sp_{hh})+c_3+c_s\}$.

The AVoI can be expressed as a piecewise linear function with respect to $c_3$:

$$AVoI = \tfrac{1}{\tau}\sum_{i\in\{ll,lh,hl,hh\}} p_i(\min_{j\in\{a,b,c,d,e\}} Cost_{n,j} - \min_{j\in\{a,b,c,d,e\}} Cost_{i,j}).$$

When $c_3 < \bar{c_3} = \min\{c_1+c_2-c_sp_{hh}, c_s(1-p_{ll})\}$, the decision maker uses flexible capacity under null information. Therefore, $AVoI$ can be expressed as $AVoI = \frac{1}{\tau}\sum\limits_{i\in\{ll,lh,hl,hh\}} p_i \max\limits_{j\in\{a,b,c,d,e\}}(Cost_{n,e} - Cost_{i,j})$, which is an increasing piecewise linear convex function with respect to $c_3$.

When $c_3 > \bar{c_3} = min\{c_1+c_2-c_sp_{hh}, c_s(1-p_{ll})\}$, the decision maker does not use flexible capacity under null information. Therefore, $AVoI$ can be expressed as $AVoI = \frac{1}{\tau}\sum\limits_{i\in\{ll,lh,hl,hh\}} p_i \max\limits_{j\in\{a,b,c,d,e\}}(B - Cost_{i,j})$, where B is constant respect to $c_3$. Hence, $AVoI$ is a non-increasing piecewise linear convex function with respect to $c_3$.

• **Proof of (1) and (3)**: We prove the first and third bullet of the Theorem by showing that $AVoI$ is in fact a decreasing function for some $c_3 > \bar{c}_3$. Let us consider the cost $\bar{c}_3 = \min\{c_1+c_2-c_s p_{hh}, c_s(1-p_{ll})\}$, where the marginal value of flexible capacity under the null information scenario is exactly 0. With partial demand visibility, we consider the case in which the decision maker observes a high low scenario ($\boldsymbol{d} = (0, 2\delta)$). With this observation, the decision maker commits to flexible capacity if the following conditions are met:

(a): $Cost_{hl,e} < Cost_{hl,d}$. This occurs if and only if $c_3 < c_1 + c_2 - \frac{(\tau-1)}{\tau} c_s p_{hh}$. This is always true when $c_3 = \bar{c}_3$:

$$c_3 = \bar{c}_3 = \min\{c_1 + c_2 - c_s p_{hh}, c_s(1 - p_{ll}) \leq c_1 + c_2 - c_s p_{hh} < c_1 + c_2 - \tfrac{(\tau-1)}{\tau} c_s p_{hh}.$$

(b): $Cost_{hl,e} < Cost_{hl,c}$. This occurs if and only if $\tau c_3 < c_s + (\tau - 1)c_s(p_{hl} + p_{lh} + p_{hh}) = c_s + (\tau - 1)c_s(1 - p_{ll})$. This is always true when $c_3 = \bar{c}_3$:

$$\tau c_3 = \tau \bar{c}_3 = \tau \min\{c_1 + c_2 - c_s p_{hh}, c_s(1 - p_{ll}) \leq \tau c_s(1 - p_{ll}) =$$
$$(\tau - 1)c_s(1 - p_{ll}) + c_s(1 - p_{ll}) < (\tau - 1)c_s(1 - p_{ll}) + c_s.$$

(c): $Cost_{hl,e} < Cost_{hl,c}$. This occurs if and only if $c_3 < c_2 + \frac{(\tau-1)}{\tau} c_s p_{lh}$. Note that if $\tau \to 1$, this is never true. In this case, when $\tau$ is very short, Region C does not exist. This occurs because visibility removes most part of uncertainty for very short commitments, and the decision maker does not use flexible capacity. Indeed, when commitment is 1, the decision maker only uses dedicated capacity to supply the visible single period. However, since $\bar{c}_3 < c_2 + c_s p_{lh}$ (strict inequality), there exist a large enough $\tau$ for which $\bar{c}_3 < c_2 + \frac{(\tau-1)}{\tau} c_s p_{lh}$. The case with $Cost_{hl,e} < Cost_{hl,c}$ is symmetric.

Therefore, when $c_3 = \bar{c}_3$, the marginal value of flexible capacity is 0 under null information and strictly greater than 0 with demand visibility (for large enough $\tau$). Hence, $AVoI = \sum_{i \in \{ll,lh,hl,hh\}} p_i \min_{j \in \{a,b,c,d,e\}} (B - Cost_{i,j})$ is a strictly decreasing function in the neighborhood $(\bar{c}_3, \bar{c}_3 + \epsilon]$ (for some $\epsilon > 0$). Hence, the maximum occurs exactly at $\bar{c}_3$.

• **Proof of (2)**: It is straightforward to show the second bullet by looking at $AVoI$ when $c_3 = \bar{c}_3$.

• **Proof of (4)**. We note that from the third bullet, AVoI is unimodal. To prove bullet (4), we show that the cost where $AVoI = AVoI_D$ (in the range of $c_3 < \bar{c}_3$) is unique (if there is any) and this cost of flexibility decreases with $\tau$. When $AVoI_D = 0$, it is easy to see that Region C is the largest $[c_1, 2c_1]$ (note that $c_1 = c_2$) and the result is straightforward. Hence, we shall focus on the case where $AVoI_D > 0$.

We first find the expression of $AVoI_D$, a symmetric dedicated network ($c_1 = c_2$, $p_{hl} = p_{lh}$). First, the null information cost is $V(\mathbf{k}_{null}) = 4\delta \min\{c_s(p_{lh} + p_{hh}), c_1\}$. The cost with partial information can be found as two independent symmetric nodes. Since $AVoI_D > 0$, commitment at each node when observed demand is high must be different from when observed demand is low. Hence, for node $i$ if demand is low, $k = 0$, if demand is high $k = 2\delta$. Therefore, $\mathbb{E}_{\mathbf{D}}\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D} | c_3 \geq 2c_1) = 2\delta(p_{ll}(\tau - 1)(c_s(2p_{hl} + 2p_{hh}) + 2p_{lh}(\tau c_1 + (\tau - 1)c_s(p_{lh} + p_{hh})) + 2p_{hh}c_1\tau)$, and $AVoI_D$ can be expressed as:

$$\begin{aligned} AVoI_D = \;& V(\mathbf{k}_{null} | c_3 \geq 2c_1) - \tfrac{1}{\tau}\mathbb{E}_{\mathbf{D}}\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D} | c_3 \geq c_1 + c_2) = \\ & 4\delta \min\{c_s(p_{lh} + p_{hh}), c_1\} - \\ & \tfrac{4\delta}{\tau}(p_{ll}(\tau - 1)(c_s(p_{hl} + p_{hh}) + p_{lh}(\tau c_1 + (\tau - 1)c_s(p_{lh} + p_{hh})) + p_{hh}c_1\tau). \end{aligned}$$

We now find the expression of $AVoI$ when $c_3 < \bar{c}_3$. Again, we focus on the case $AVoI > 0$, hence commitment under partial demand information must be different than under null information. Under null information: $V(\mathbf{k}_{null}) = 2\delta(c_3 + c_s p_{hh})$. Under partial information we have that $k_3 > 0$ when $\boldsymbol{d} = (d_l, d_h)$ and $\boldsymbol{d} = (d_h, d_l)$ (due to the condition $c_3 < \bar{c}_3$ and that $\tau$ is large enough, as shown in the proof of the first and third bullet above). We only need to consider three cases:

$$\begin{aligned}
\text{(a)}: \quad \mathbb{E}_{\mathbf{D}}\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D}) = \quad & 2\delta p_{ll}(\tau - 1)c_s(2p_{hl} + 2p_{hh}) + \\
& 4\delta p_{lh}\{(\tau - 1)(c_3 + c_s p_{hh}) + c_3\} + \\
& 4\delta p_{hh}\tau c_1.
\end{aligned}$$

$$\begin{aligned}
\text{(b)}: \quad \mathbb{E}_{\mathbf{D}}\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D}) = \quad & 2\delta p_{ll}\{(\tau - 1)(c_3 + c_s p_{hh}) + c_3\} + \\
& 4\delta p_{lh}\{(\tau - 1)(c_3 + c_s p_{hh}) + c_3\} + \\
& 4\delta p_{hh}\tau c_1.
\end{aligned}$$

$$\begin{aligned}
\text{(c)}: \quad \mathbb{E}_{\mathbf{D}}\text{Cost}_{Total}(\mathbf{k}^*(\mathbf{D}), \mathbf{D}) = \quad & 2\delta p_{ll}(\tau - 1)c_s(2p_{hl} + 2p_{hh}) + \\
& 4\delta p_{lh}\{(\tau - 1)(c_3 + c_s p_{hh}) + c_3\} + \\
& 2\delta p_{hh}\{(\tau - 1)(c_3 + c_s p_{hh}) + c_3 + c_s\}.
\end{aligned}$$

We show that in any of these cases, the cost $c_3$ where $AVoI_D - AVoI = 0$ is decreasing in $\tau$. Note that by dividing the entire equation by $2\delta$, we cancel out the impact of variability.

For case (a):

$$0 = \frac{1}{2\delta}(AVoI_D - AVoI) = \quad 2\min\{c_s(p_{lh}+p_{hh}), c_1\}+$$
$$-p_{ll}\tfrac{\tau-1}{\tau}c_s(2p_{hl}+2p_{hh}) - 2p_{lh}(c_1 + \tfrac{\tau-1}{\tau}c_s(p_{lh}+p_{hh})) - 2p_{hh}c_1+$$
$$-(c_3 + c_s p_{hh})+$$
$$+p_{ll}\tfrac{\tau-1}{\tau}c_s(2p_{hl}+2p_{hh}) + 2p_{lh}\{\tfrac{\tau-1}{\tau}(c_s p_{hh}) + c_3\} + 2p_{hh}c_1.$$

We can obtain an expression for $c_3$ that solves the previous equation, which is decreasing with $\tau$ and independent of $\delta$:

$$c_3 = \tfrac{2}{1-2p_{lh}}(\min\{c_s(p_{lh}+p_{hh}), c_1\} - \tfrac{1}{2}c_s p_{hh} - p_{lh}c_1 - c_s p_{lh}^2\tfrac{\tau-1}{\tau}).$$

For case (b):

$$0 = \frac{1}{2\delta}(AVoI_D - AVoI) = \quad 2\min\{c_s(p_{lh}+p_{hh}), c_1\}+$$
$$-p_{ll}\tfrac{\tau-1}{\tau}c_s(2p_{hl}+2p_{hh}) - 2p_{lh}(c_1 + \tfrac{\tau-1}{\tau}c_s(p_{lh}+p_{hh})) - 2p_{hh}c_1+$$
$$-(c_3 + c_s p_{hh})+$$
$$+p_{ll}(\tfrac{\tau-1}{\tau}c_s p_{hh} + c_3) + 2p_{lh}\{\tfrac{\tau-1}{\tau}c_s p_{hh} + c_3\} + 2p_{hh}c_1.$$

We can obtain an expression for $c_3$ that solves the previous equation, which is decreasing with $\tau$ and independent of $\delta$:

$$c_3 = \tfrac{2}{1-2p_{lh}-p_{ll}}(\min\{c_s(p_{lh}+p_{hh}), c_1\} - \tfrac{1}{2}c_s p_{hh} - p_{lh}c_1 - c_s p_{lh}^2\tfrac{\tau-1}{\tau} - p_{ll}\tfrac{\tau-1}{\tau}c_s(p_{hl}+\tfrac{1}{2}p_{hh})).$$

For case (c): The result is similarly obtained.

□

## B.2. The Impact of Demand Visibility on Dedicated Commitments

Recall from Section 3.4.3 that committed dedicated capacity for node 1 $k_1^*(\boldsymbol{d})$ depends on demand $d_2$ observed at the other node. In Proposition 9, we compare the bounding regions to show that the commitment to dedicated capacity of a node decreases when observed demand of the two nodes is largely different with high observed demand at one node and low observed demand at the other node.

**Proposition 9.** *[**The Impact of Demand Visibility on Dedicated Commitments**] If the observed demand of one node is high (enough) and that of the other node is low (enough), as their difference grows, the decision maker reduces the dedicated capacity of both nodes. Specifically:*

*(a) If observed demand at node* 1 *(2) is high enough, committed dedicated capacity for node* 1 *(2) is lower when observed demand at node 2 (1) is low than when observed demand at node 2 (1) it is high. ($\psi_2^I < \psi_2^{III}$ and $\psi_1^{II} < \psi_1^{III}$.)*

*(b) If observed demand at node 1 (2) is low enough, committed dedicated capacity for node 1 (2) is lower when observed demand at node 2 (1) is high than when observed demand at node 2 (1) it is low. ($\psi_1^I < \psi_1^0$ and $\psi_2^{II} < \psi_2^0$.)*

Proposition 9 contrasts with the results without flexible capacity in Section 3.2, where committed dedicated capacity at node 1 is independent of the demand observed at node 2; i.e., it does not depend on whether $d_2$ is low or high. The simultaneous availability of flexible capacity and demand visibility changes the contracting policies for dedicated contracts with commitment.

### B.3. The case with $\mathbf{c} \prec \frac{c_s}{\tau}$

The value of $\boldsymbol{\psi}^i$ depends on the costs, the commitment horizon and the demand distribution. When the commitment horizon is very short or the cost of committed capacity is very inexpensive (or, conversely, spot market costs very expensive), the decision maker has an incentive to fulfill all the observed demand using committed capacity. Under these conditions, the range of adaptation is unbounded and, therefore, the higher is the demand observed, the higher is the committed capacity. Proposition 10 shows that when $\mathbf{c} \prec \frac{c_s}{\tau}$, the cost of the contract is so low (or commitment is so short) that the optimal committed capacity always fully satisfies the observed demand.

**Proposition 10.** *A.3*

- *If $c_1 < \frac{c_s}{\tau}$, then $\hat{\Omega}_I = \hat{\Omega}_{III} = \emptyset$ and all the observed next period demand of node 1 is fully satisfied using committed capacity $k_1$ or flexible capacity $k_3$.*
- *If $c_2 < \frac{c_s}{\tau}$, then $\hat{\Omega}_{II} = \hat{\Omega}_{III} = \emptyset$ and all the observed next period demand of node 2 is fully satisfied using committed capacity $k_2$ or flexible capacity $k_3$.*

Under the scenario where $c_1 \leq c_2 < \frac{c_s}{\tau}$, from Proposition 10, we none of the bounding regions $\hat{\Omega}_I$, $\hat{\Omega}_{II}$ or $\hat{\Omega}_{III}$ to exist. In that case, $k_1^*(\mathbf{d})$ and $k_2^*(\mathbf{d})$ are unbounded. Hence, dedicated capacities monotonically increase as observed demand $(d_1, d_2)$ increases. Note that, while Proposition 10 shows conditions for which the bounding regions $\hat{\Omega}_I, \hat{\Omega}_{II}$ and $\hat{\Omega}_{III}$ vanish, the region $\Omega_0$ is never empty. Therefore, the committed capacities are never arbitrarily small.

## B.4. Additional Numerical Experiments

**Definition 7.**

- *We define Value of Pooling ($VoP$) as the expected cost reduction of using the flexible contract relative to the no network case:*

$$VoP = 100\left(1 - \frac{\mathbb{E}_{\mathbf{D}}[Cost_{Total}(\mathbf{k}^*(\mathbf{D}),\mathbf{D})|(c_1,c_2,c_3)]}{\mathbb{E}_{\mathbf{D}}[Cost_{Total}(\mathbf{k}^*(\mathbf{D}),\mathbf{D})|(c_1,c_2,\infty)]}\right). \tag{B.5}$$

- *We define Value of Pooling with No Information ($VoP_{Null}$) as the $VoI$ when $\tau \to \infty$.*

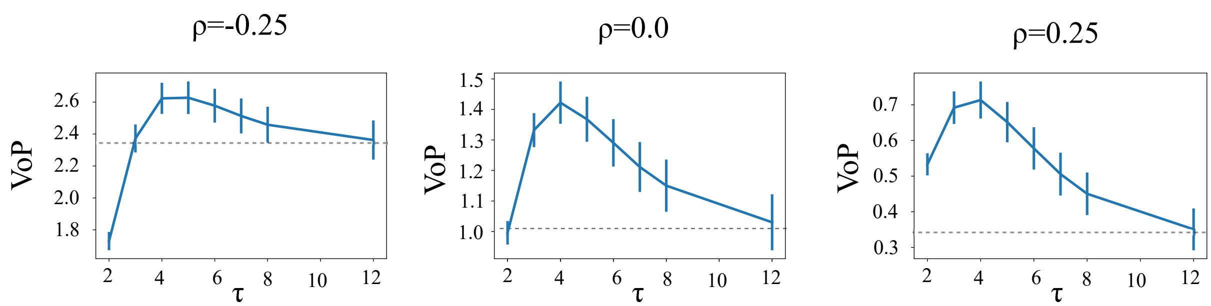


Figure B.1. VoP for different commitment horizons and correlation coefficient with bivariate Gamma distribution and $\mathbf{c} = (0.3, 0.3, 0.45)$.

Similarly to VoI, there always exists a region where the value of flexible capacity with partial visibility is higher than without partial visibility.

APPENDIX C

# Appendix from Chapter 4

## C.1. Including All Customers

To ease our exposition, we have only considered the customers that scheduled the home delivery within 28 days. We can easily incorporate the rest of the customer by adding an additional layer in our sequential model. We consider two different customer segments: (1) Customers that choose a delivery date within 28 days as described in the main text. (2) Customers that planned to choose a date beyond the 28th day after purchase. The modified version of the model incorporates additional parameters to estimate the probability that a customer belongs to either segment. Hence, the sequential probability equations 4.2 would have modified initial conditions that incorporate the probability that customers choose a delivery date within 28 days.

## C.2. Assumption 1: Home Delivery Teams have fixed delivery capacity

In this section we use the data from the furniture company to support the assumption that the average capacity per team is constant.

In Figure C.1-(a) we show, for each day and region, the total number of orders delivered, and the number of delivery teams assigned. The redline shows the regression between Orders Delivered (OD) and Number of Teams (NT). Looking at the increase of OD per each unit of NT allows us to estimate the delivery capacity. On average, each additional team increased OD by 4.99. We relate this number to the per-team capacity because it is the amount of orders each team can contribute to the total OD. We note that the slope of the regression approximately coincides the 66th percentile of the distribution of Orders Delivered per Delivery Team ($\frac{\text{OD}}{\text{NT}}$), as shown in Figure C.1-(b).

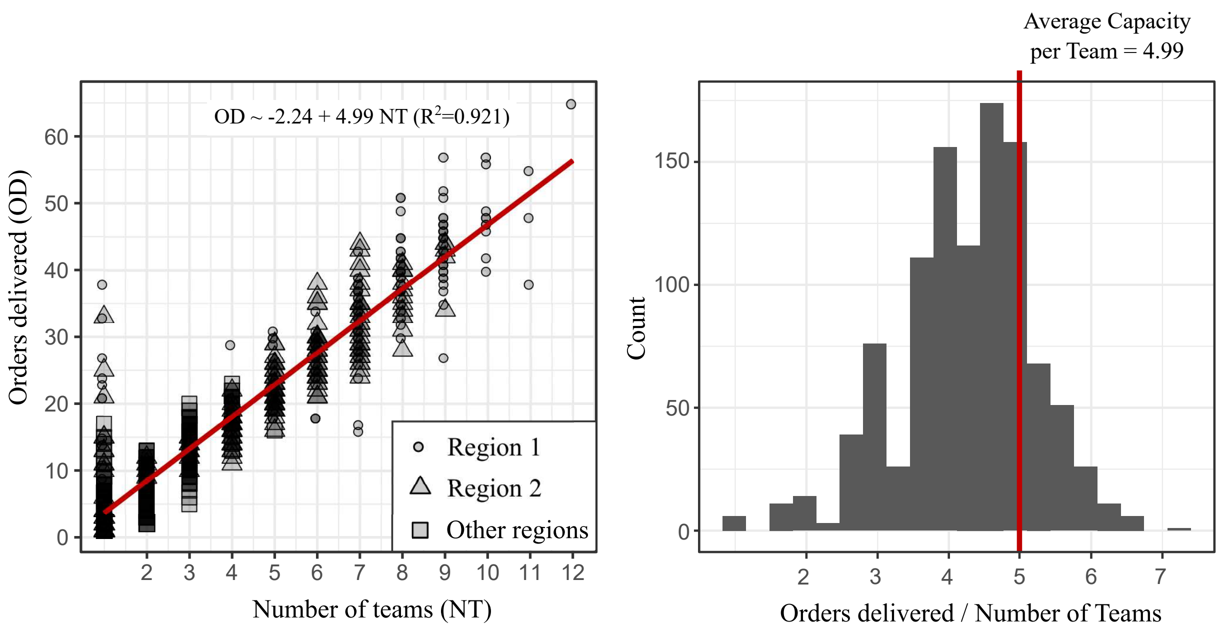


Figure C.1. Estimation of average capacity per team. Figure (a) shows the numbers of orders delivered and the number of teams dedicated to each region in all days in the season 2016-2017. Figure (b) shows the histogram of the orders per team delivered each day. The regression and the histogram in (b) do not incorporate the data points with only one team.

We allow a different team capacity for each region. We denote $OD_{r,d}$ and $NT_{r,d}$ the Orders Delivered and the Number of Teams in region $r$ on day $d$, respectively. We then regress the model in Equation C.1 and report the results in Table C.1. Observe that our model, which assumes a constant per-team delivery capacity, explains 92% of the variability in OD.

| Variables | Coefficient |
|---|---|
| Intercept | -1.42 (0.48)*** |
| Region 1 x OD | 4.93 (0.09)*** |
| Region 2 x OD | 4.71 (0.1)*** |
| Region 3 x OD | 4.93 (0.17)*** |
| Region 4 x OD | 5.06 (0.19)*** |
| Region 5 x OD | 4.28 (0.24)*** |
| Region 6 x OD | 4.41 (0.31)*** |
| Region 7 x OD | 3.68 (0.35)*** |
| Region 8 x OD | 3.65 (0.29)*** |
| Region 9 x OD | 3.71 (0.29)*** |
| N Obs. | 1053 |
| N Pars. | 10 |
| $R^2$ | 0.9266 |

Table C.1. Results of regression of Equation C.1. Parenthesis show cluster robust standard errors (***p<0.001.)

| Region 1 | Region 2 | Region 3 | Region 4 | Region 5 | Region 6 | Region 7 | Region 8 | Region 9 |
|---|---|---|---|---|---|---|---|---|
| 5.00 | 4.66 | 4.72 | 5.00 | 4.00 | 4.00 | 3.50 | 3.50 | 3.50 |

Table C.2. 66th percentile of the ratio $\frac{OD}{NT}$ for each region.

$$OD_{r,d} = \beta_0 + \beta_r NT_{r,d} + \epsilon_{r,d}. \tag{C.1}$$

The values of the coefficients $\beta_r$ approximately coincide with the 66th percentile of the observed ratio $\frac{OD}{NT}$ of each region, shown in Table C.2. In our study of Chapter 4 we use the values in Table C.2 as the per-team capacity in each region. As a robustness check, we also replicate the analysis of Chapter 4 with percentiles 50th, 60th, 80th and 85th. In all cases, we obtain very similar results, showing that our results are robust to the per-team capacity estimate.

### C.3. Assumption 2: The Impact of Lost Sales Is Negligible

In this section we study the impact of day availability on lost sales. Instead of considering sales at the individual order level, we consider order volume, i.e.; the aggregated total of orders entered in the system in a given time period (similarly to [**74, 82**]). The variable that we use to capture the impact of day availability is the number of days with no remaining delivery capacity within the first 28 days, we call this variable Day Occupancy (DO). We study whether instances with large DO are linked with lower number of orders (higher lost sales).

The DO changes within a day because delivery capacity is depleted as new orders are scheduled for delivery. Customers that arrive later in the store observe higher DO than customer that arrive earlier. To control for changes in DO and sales within a day, we split each day in three time buckets: From store opening time to 12pm, from 12pm to 4pm, and from 4pm to store closing time. In Table C.3 we show the average change of DO within a time bucket for the four largest regions. Observe that within a time bucket, DO change on average less than 5%. In other words, if region $r$ on day $d$ at time bucket $t$ has an average DO $\overline{DO}_{r,d,t}$, customers purchasing in that same region, day and time bucket observe a DO within the range $(1 \pm 0.05)\overline{DO}_{r,d,t}$. In Table C.4 we show summary statistics of the total number of orders and DO of each region and time bucket. Note that, the Coefficient of Variation across days and regions (sixth column of Table C.3), which is around 0.54, is large compared to the change of DO within a time bucket (third column of Table C.3).

To study the impact of DO on volume of orders we use a logarithmic regression model as in [**61**] , [**74**] and [**82**]. We denote $Vol_{r,d,t}$ and $DO_{r,d,t}$ as the volume of orders and the

| Region | Time bucket | Av. change in DO |
|---|---|---|
| Region 1 | Opening - 12pm | 3.3% |
| | 12pm - 4pm | 2.5% |
| | 4m - Closing | 2.5% |
| Region 2 | Opening - 12pm | 1.1% |
| | 12pm-4pm | 4.1% |
| | 4m - Closing | 2.6% |
| Region 3 | Opening - 12pm | 1.9% |
| | 12pm-4pm | 3.8% |
| | 4m - Closing | 4.9% |
| Region4 | Opening - 12pm | 2.5% |
| | 12pm-4pm | 5.1% |
| | 4m - Closing | 1.6% |

Table C.3. Average change within each grade bucket.

| Region | Time bucket | Mean volume orders * | CV volume orders | Mean DO | CV DO |
|---|---|---|---|---|---|
| Region 1 | Opening - 12pm | 0.69 | 0.73 | 2.06 | 0.56 |
| | 12pm - 4pm | 0.97 | 0.77 | 2.11 | 0.55 |
| | 4m - Closing | 1.00 | 0.83 | 2.20 | 0.57 |
| Region 2 | Opening - 12pm | 0.51 | 0.77 | 2.02 | 0.55 |
| | 12pm - 4pm | 0.66 | 0.77 | 2.05 | 0.56 |
| | 4m - Closing | 0.72 | 0.85 | 2.12 | 0.56 |
| Region 3 | Opening - 12pm | 0.25 | 0.86 | 2.28 | 0.53 |
| | 12pm - 4pm | 0.34 | 0.85 | 2.35 | 0.52 |
| | 4m - Closing | 0.35 | 0.96 | 2.43 | 0.52 |
| Region 4 | Opening - 12pm | 0.24 | 0.83 | 1.95 | 0.54 |
| | 12pm - 4pm | 0.32 | 0.81 | 2.00 | 0.55 |
| | 4m - Closing | 0.32 | 0.94 | 2.08 | 0.55 |

Table C.4. Summary statistics of regions and time buckets.(*Normalized by the maximum value in the column.)

average number of days without remaining capacity in region $r$, day $d$ and time bucket $t$, respectively. In our regression we also include categorial variables to control for region ($Region_r$), time bucket ($Time_t$), day of week ($DOW_d$) and a biweekly seasonality term ($Season_d$) as in Equation 4.4. In Equation C.2 we specify our model.

(C.2)

$$\text{Model } DO^{1-28}\text{:}\quad \log(1+Vol_{r,d,t}) = \beta_0 + \beta DO_{r,d,t} + \alpha_r^R Region_r + \alpha_d^S Season_d + \alpha_d^{DOW} DOW_d + \alpha_{d,t}^{DOWXT} DOW_d \text{x} Time_t + \epsilon_{r,d,t}.$$

We additionally consider two more specifications. The second specification in Equation C.3 only consider the DO within the first week after purchase, we denote this variable $DO^{1-6}$. The third specification in Equation C.4 consider the DO within the second week after purchase, we denote this variable $DO^{7-13}$.

(C.3)

$$\text{Model } DO^{1-6}\text{:}\quad \log(1+Vol_{r,d,t}) = \beta_0 + \beta DO_{r,d,t}^{1-6} + \alpha_r^R Region_r + \alpha_d^S Season_d + \alpha_d^{DOW} DOW_d + \alpha_{d,t}^{DOWXT} DOW_d \text{x} Time_t + \epsilon_{r,d,t}.$$

(C.4)

$$\text{Model } DO^{7-13}\text{:}\quad \log(1+Vol_{r,d,t}) = \beta_0 + \beta DO_{r,d,t}^{7-13} + \alpha_r^R Region_r + \alpha_d^S Season_d + \alpha_d^{DOW} DOW_d + \alpha_{d,t}^{DOWXT} DOW_d \text{x} Time_t + \epsilon_{r,d,t}.$$

We focus our analysis on the impact of $DO_{r,d,t}$ through $\boldsymbol{\beta}$. To that aim, in Table C.5 we present the regression results of Equations C.2-C.4.

We find that the impact of $DO$ is not significant in models Model $DO^{1-28}$, and Model $DO^{6-7}$. There is a significant positive effect of $DO^{7-13}$, suggesting that the higher is DO, the larger is the volume of orders entered in the system. However, the regressions in Table C.5 suffer from endogeneity due to simultaneity bias: Observations with high volume of orders registered tend to have more DO because delivery capacity is depleted. To deal with the endogenous relation we use a two-step linear regression approach with

| Variables | Model $DO^{1-28}$ | Model $DO^{1-6}$ | Model $DO^{7-13}$ |
|---|---|---|---|
| Intercept | 1.181 (0.050)*** | 1.216 (0.052)*** | 1.199 (0.048)*** |
| $DO$ | 0.0075 (0.006) | | |
| $DO^{1-6}$ | | -0.006 (0.008) | |
| $DO^{7-13}$ | | | 0.027 (0.0011)** |
| $Region$ | yes | yes | yes |
| $Season$ | yes | yes | yes |
| $DOW$ | yes | yes | yes |
| $DOW$ x $Time$ | yes | yes | yes |
| NoPars | 68 | 68 | 68 |
| No Obs | 7074 | 7074 | 7074 |
| $R^2$ | 0.6018 | 0.6018 | 0.6022 |

Table C.5. Results of regression of Equations C.2-C.4. Parenthesis show cluster robust standard errors (**p<0.01 ***p<0.001.)

instrumental variables [**140**]. Similarly to the procedure followed by [**139**], [**118**], [**35**], [**21**], we use lagged versions of the endogenous variables. Following the arguments of [**21**], lagged versions of the DO variables are valid instruments: First, customers in time bucket $t$ are not impacted by the DO of the previous time bucket at $t-1$ (if the time bucket $t$ is the first of day $d$, we define time bucket $t-1$ as the last time bucket of day $d-1$). Second, a Durbin-Watson test for serial correlation showed that volume of orders is not autocorrelated. Third, DO of time bucket $t-1$ is highly positively correlated with time bucket $t$ as shown below.

In Table C.6 we show the results of the first step, which regresses the model in Equation C.5 for $DO^{7-13}$.

$$\begin{aligned}\text{Model 1-step:}\quad DO^{7-13}_{r,d,t} =& \nu^R_r Region_r + \nu^S_d Season_d + \\ & \nu^{DOW}_d DOW_d + \nu^{DOWXT}_{d,t} DOW_d \text{x} Time_t + u_{r,d,t}.\end{aligned} \tag{C.5}$$

| Variables | Model 1-Step |
|---|---|
| Intercept | 0.041 (0.018) |
| Lagged $DO^{7-13}$ | 0.918 (0.007)*** |
| *Region* | yes |
| *Season* | yes |
| *DOW* | yes |
| *DOW* x *Time* | yes |
| NoPars | 68 |
| No Obs | 7074 |
| $R^2$ | 0.8901 |

Table C.6. Results of regression of Equations C.5. Parenthesis show cluster robust standard errors (**p<0.01 ***p<0.001.)

We then use the fitted values of the model of Equation C.5, $\widehat{DO}^{7-13}_{r,d,t}$, in substitution of $DO^{7-13}_{r,d,t}$ in Equation C.4. We show the results in Table C.7. Observe that once we deal with simultaneity bias, the impact of DO on volume of orders vanishes. Replicating the same steps for the specification in Equations C.2 and C.3 we obtain there is no statically significant impact of DO on volume orders. These results support our intuition that the impact of day availability on order volume due to lost sales is negligible, i.e.; the second assumption in Section 4.3.2.

## C.4. Assumption 3: From among all offered dates, customers choose the earliest available date.

In this section we present results that support the third assumption in Section 4.3.2: From among all offered dates, customers choose the earliest available date.

If the assumption is true the next logic follows: consider that from among all the dates offered to a customer, the customer chooses day $d_j$. If $d_j$ is chosen because it is the earliest day offered that is available to the customer, offering later dates would not impact the

| Variables | Model 2-Step |
| --- | --- |
| Intercept | 1.200 (0.047)*** |
| Fitted Values ($\widehat{DO}^{7-13}_{r,d,t}$) | 0.004 (0.010) |
| *Region* | yes |
| *Season* | yes |
| *DOW* | yes |
| *DOW* x *Time* | yes |
| NoPars | 68 |
| No Obs | 7074 |
| $R^2$ | 0.6018 |

Table C.7. Results of regression of the model in Equation C.4 using fitted variables. Parenthesis show cluster robust standard errors (**p<0.01 ***p<0.001.)

customer choice. On the other hand, offering additional earlier dates than $d_j$ can make the customer choose an earlier date, if that date is available for the customer. Therefore, offering day $d_e$ that comes earlier than $d_j$ decreases the probability that $d_j$ is chosen. On the other hand, offering day $d_l$ that comes later that $d_j$ does not impact the customer decision. Hence, if our assumption holds, offering day $\tilde{d}$ impacts more strongly customers that were to choose later dates (later than $\tilde{d}$) than customer that were to choose earlier dates (earlier than $\tilde{d}$).

We aim to know if availability of earlier dates reduces the odds that customers choose later dates. In particular we aim to show that if a day $\tilde{d}$ in the schedule is offered, the odds that customer chooses a later day $d_l$ (i.e., $d_l > \tilde{d}$) against an earlier date $d_e$ (i.e., $d_e < \tilde{d}$) decreases.

To support the third assumption in Section 4.3.2 we first describe an idealized experiment (not feasible to be performed) that would allow us to test whether the assumption holds. We then show how we can use natural experiments in our data to replicate the experiment.

**Idealized Experiment**

In an ideal setting we could perform two identical experiments in which we offer the same customer two different set of dates for home delivery. By repeating the experiment with several customers, we would be able to study how customers' choice differed when choosing from among one set versus the other.

We define the *schedule profile* as the set of days with remaining delivery capacity that can be potentially offered to a customer. Focusing on the first 28 days, a schedule profile can be represented as a 28-length binary vector where each element $j$ is 0 or 1 depending on whether day $d_j$ has remaining delivery capacity (1 if the day can be offered and 0 otherwise).

To study whether offering or not offering $d_k$ has an impact on customer choice, we would consider two different sets $v_0$ and $v_1$ that are identical but only differ on element $k$, which is 0 in $v_0$, and 1 in $v_1$. In our ideal setting, we would like to test the same customer in two different scenarios: A scenario $S_0$ where days are offered according to schedule profile $v_0$, and a scenario $S_1$ where days are offered according to schedule profile $v_1$. In this idealized experiment, we would study whether the odds that a customer chooses later days $d_l$ such that $d_l > d_k$, against earlier dates $d_e$ such that $d_e < d_k$, is lower in $S_1$ than in scenario $S_0$. In other words, if customer chooses day $d_J$[1] we would like to test whether Equation C.6 is true.

[1]Note $d_J$ has capital index $J$ because the choice of the customer is a random variable whose realizations are the days with remaining delivery capacity offered.

$$\text{(C.6)} \qquad \text{Odds } S_1 := \frac{P[d_J > d_k|S_1]}{P[d_J < d_k|S_1]} < \frac{P[d_J > d_k|S_0]}{P[d_J < d_k|S_0]} =: \text{Odds } S_0.$$

By repeating the experiment with several different customers, we could test if Equation C.6 holds. If Equation C.6 holds, the experiment would support the assumption that availability of day $k$ has attracted more customers that were to chose later dates than customers that were to chose earlier dates, reducing the odds $\frac{P[d_J > d_k|S_1]}{P[d_J < d_k|S_1]}$ with respect to scenario $S_0$. Then, if we find that Equation C.6 holds for all $k$, the set of experiments would support the logic that follows from our third assumption in Section 4.3.2.

Unfortunately, such an experiment is impossible to perform. However, in our data, the days offered to customers are subject to delivery capacity constraints and, consequently, customers arriving in different time and region are likely to be offered different sets of delivery dates. We explore the transactional data from the furniture company to find natural experiments similar to our idealized setting.

**Natural Experiments**

To replicate the idealized experiment, we find in our data instances where two different customers where offered "similar" sets of delivery dates.

Before presenting our results we first describe how we can compare two different schedule profiles, i.e.; we define what two "similar" sets are. We say that two schedule profiles $v_i$ and $v_j$ are *k-comparable* if: (1) they only differ in the kth element, (2) there is at least a day offered earlier than $d_k$, and (3) there is at least a day offered later than $d_k$. The second and third conditions make sure that there are alternative days earlier and later than $d_k$, otherwise, probabilities $P[d_J > d_k|S_1] = 0$ or $P[d_J < d_k|S_1] = 0$. In Figure C.2, we show examples of schedule profiles that are $k$-comparable and examples that are not $k$-comparable. For a $k$-comparable pair, we denote $v_0$ ($v_1$) as the schedule profile with a 0 (1) in element $k$.

| 3-Comparable | No Comparable | No Comparable |
|---|---|---|
| [0, 1, 0, 1, 1, 0, 0]<br>[0, 1, 1, 1, 1, 0, 0] | [0, 1, 0, 1, 1, 0, 1]<br>[0, 1, 1, 1, 1, 1, 1] | [0, 1, 1, 0, 1, 0, 0]<br>[1, 1, 1, 0, 1, 0, 0] |
| **6-Comparable** | **No Comparable** | **No Comparable** |
| [0, 1, 1, 0, 1, 1, 1]<br>[0, 1, 1, 0, 1, 0, 1] | [0, 0, 0, 1, 1, 0, 1]<br>[0, 0, 1, 1, 1, 0, 1] | [0, 1, 0, 0, 0, 0, 0]<br>[0, 1, 1, 0, 0, 0, 0] |

Figure C.2. Examples of K-comparable vectors and non comparable vectors of length 7.

To remove the effect of day of week, we only compare schedule profiles of customers that made a purchase on the same day of week. We look at pairs of comparable profiles that have at least 50 observations each. In other words, given that two schedule profiles $v_i$ and $v_j$ are $k$-comparable, we consider the experiment valid only if $v_i$ and $v_j$ have been offered at least to 50 customers, respectively. We define the empirical ratios similarly as in Equation C.6 using the empirical proportions:$\hat{P}[d_J > d_k|S_1]$ ($\hat{P}[d_J < d_k|S_1]$) is the proportion of customers that were offered $v_1$ and chose a delivery date later (earlier) than $d_k$; similarly, $\hat{P}[d_J > d_k|S_0]$ ($\hat{P}[d_J < d_k|S_0]$) is the proportion of customers that were offered $v_0$ and chose a delivery date later (earlier) than $d_k$.

In our data there are 28 cases that satisfy the experimental conditions, which we summarize in Table C.8. The first column shows the schedule profile of each experiment, where the 'X' shows the position from where $v_0$ and $v_1$ differ (i.e.; the $k$th position). In 27 experiments out of the 28 we find that the odds ratio when $v_1$ is offered, $\frac{\hat{P}[d_J > d_k|S_1]}{\hat{P}[d_J < d_k|S_1]}$ (10th column), is lower than the odds ratio when $v_0$ is offered, $\frac{\hat{P}[d_J > d_k | S_0]}{\hat{P}[d_J < d_k | S_0]}$ (11th column). In other words, in 27 out of 28 cases, offering day $\tilde{d}$ has more strongly impacted customers that were to choose later dates than customer that were to choose earlier dates.

While more rigourous analysis is needed to formally conclude that availability of earlier dates reduces the odds that customers choose later dates, the aim of this section is to present a set preliminary results that suggest that our third assumption is reasonable. While statistical analysis must be performed, 27 out of 28 natural experiments found in our data seem to confirm the intuition of the third assumption. To formalize this section we suggest testing the hypothesis behind the assumption using formal econometric tools such as regression and/or matching.

| Sch. Profile | $k$ | DoW | $N_1$ | $N_0$ | $\hat{P}_{e,1}$ | $\hat{P}_{e,0}$ | $\hat{P}_{l,1}$ | $\hat{P}_{l,0}$ | $\widehat{\text{Odds}}\ S_1$ | $\widehat{\text{Odds}}\ S_0$ | $\widehat{\text{Odds}}\ S_1 < \widehat{\text{Odds}}\ S_0$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 001X100111 | 4 | Mon | 1120 | 208 | 0.04 | 0.01 | 0.85 | 0.94 | 19.14 | 97.50 | ✓ |
| 0011X00111 | 5 | Mon | 1120 | 219 | 0.15 | 0.11 | 0.76 | 0.81 | 5.23 | 7.08 | ✓ |
| 0011100X11 | 8 | Mon | 1120 | 77 | 0.24 | 0.18 | 0.62 | 0.82 | 2.60 | 4.50 | ✓ |
| 0010100X11 | 8 | Mon | 208 | 71 | 0.21 | 0.06 | 0.66 | 0.92 | 3.11 | 16.25 | ✓ |
| 001X100011 | 4 | Mon | 77 | 71 | 0.03 | 0.01 | 0.90 | 0.96 | 34.50 | 68.00 | ✓ |
| 001100X111 | 7 | Tue | 991 | 107 | 0.12 | 0.12 | 0.71 | 0.87 | 5.87 | 7.15 | ✓ |
| 000100X111 | 7 | Tue | 192 | 50 | 0.05 | 0.04 | 0.82 | 0.94 | 17.44 | 23.50 | ✓ |
| 001X000111 | 4 | Tue | 107 | 59 | 0.04 | 0.02 | 0.88 | 0.97 | 23.50 | 57.00 | ✓ |
| 001000X111 | 7 | Tue | 249 | 59 | 0.16 | 0.03 | 0.71 | 0.81 | 4.56 | 24.00 | ✓ |
| 001X001111 | 4 | Tue | 991 | 249 | 0.06 | 0.10 | 0.88 | 0.85 | 15.63 | 8.12 | X |
| 001001X111 | 7 | Wed | 981 | 69 | 0.19 | 0.04 | 0.68 | 0.91 | 3.68 | 21.00 | ✓ |
| 00100X1111 | 6 | Wed | 981 | 156 | 0.01 | 0.00 | 0.81 | 0.98 | 159.80 | Inf | ✓ |
| 00001X1110 | 6 | Thu | 1261 | 139 | 0.07 | 0.03 | 0.80 | 0.93 | 12.02 | 32.25 | ✓ |
| 0001X11100 | 5 | Fri | 1577 | 179 | 0.01 | 0.01 | 0.92 | 0.98 | 161.78 | 175.00 | ✓ |
| 00011X1100 | 6 | Fri | 1577 | 87 | 0.08 | 0.07 | 0.77 | 0.90 | 10.04 | 13.00 | ✓ |
| 0001X11001 | 5 | Sat | 454 | 75 | 0.04 | 0.04 | 0.83 | 0.91 | 19.89 | 22.67 | ✓ |
| 000111X001 | 7 | Sat | 454 | 124 | 0.31 | 0.23 | 0.56 | 0.69 | 1.79 | 2.93 | ✓ |
| 00111X1001 | 6 | Sat | 2085 | 119 | 0.17 | 0.07 | 0.69 | 0.89 | 4.08 | 13.25 | ✓ |
| 0011X11001 | 5 | Sat | 2085 | 334 | 0.06 | 0.01 | 0.83 | 0.96 | 14.79 | 80.50 | ✓ |
| 00101X1001 | 6 | Sat | 252 | 72 | 0.19 | 0.17 | 0.68 | 0.93 | 3.49 | 5.00 | ✓ |
| 0010X11001 | 5 | Sat | 252 | 93 | 0.01 | 0.00 | 0.81 | 0.91 | 101.50 | Inf | ✓ |
| 001X110011 | 4 | Sun | 1000 | 228 | 0.03 | 0.01 | 0.87 | 0.97 | 33.42 | 73.67 | ✓ |
| 0011X10011 | 5 | Sun | 1000 | 66 | 0.13 | 0.06 | 0.76 | 0.88 | 5.81 | 14.50 | ✓ |
| 00111X0011 | 6 | Sun | 1000 | 68 | 0.24 | 0.16 | 0.66 | 0.79 | 2.76 | 4.91 | ✓ |
| 00111100X1 | 9 | Sun | 1000 | 63 | 0.35 | 0.16 | 0.55 | 0.79 | 1.58 | 5.00 | ✓ |
| 0010X10011 | 5 | Sun | 228 | 58 | 0.03 | 0.00 | 0.85 | 0.95 | 27.57 | Inf | ✓ |
| 001X010011 | 4 | Sun | 66 | 58 | 0.02 | 0.00 | 0.94 | 1.00 | 62.00 | Inf | ✓ |
| 0001X10011 | 5 | Sun | 211 | 89 | 0.09 | 0.06 | 0.77 | 0.90 | 9.06 | 16.00 | ✓ |

Table C.8. Summary of the natural experiments. Abbreviations: Sch. Profile = Schedule profile, DoW = Day of week, $N_1$ ($N_0$) = Number of customers that were offered $v_1$ ($v_0$), $\hat{P}_{e,1} = \hat{P}[d_J < d_k|S_1]$, $\hat{P}_{e,0} = \hat{P}[d_J < d_k|S_0]$, $\hat{P}_{l,1} = \hat{P}[d_J > d_k|S_1]$, $\hat{P}_{l,0} = \hat{P}[d_J > d_k|S_0]$.